



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 96101

TO: Patricia Duffy
Location: cm1/8d05/8e12
Art Unit: 1645
Thursday, June 19, 2003
Case Serial Number: 674277

From: Paul Schulwitz
Location: Biotech-Chem Library
CM1-6B06
Phone: 305-1954

paul.schulwitz@uspto.gov

Search Notes

Examiner Duffy,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(703)305-1954

THIS PAGE BLANK (USPTO)

96101

From: Duffy, Patricia
Sent: Sunday, June 08, 2003 10:49 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search 09/674,277

Importance: High

RECEIVED

JUN - 9 2003

(STIC)

In re: 09/674,277

Please search SEQ ID NO:1 and 2.

Please include and interference and commercial database search.

Please print out top 50 hits in each category.

Thanks,

Patricia A. Duffy

CM1-8D05

AU 1645

703-305-7555

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/10
Date Completed: 6/19
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 15:59:59 ; Search time 2128.45 Seconds
(without alignments)
16148.066 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcaggagatggaagaaaa.....ttttactttttctgtgag 1181

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vir.*
- 38: em_sy.*
- 39: em_htg_hum.*
- 40: em_htg_mus.*
- 41: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1181	100.0	1181	6	AX011298	Sequence
c 2	1166.2	98.7	7395	1	EC74PO157	Y11275 E.coli 7.4
c 3	1166.2	98.7	7395	6	AX191728	Sequence
c 4	1166.2	-98.7	92077	1	AF074613	AF074613 Escherich
c 5	1166.2	98.7	92077	6	AX191727	Sequence
c 6	1166.2	98.7	92721	1	AB011549	AB011549 Escherich
c 7	1166.2	98.7	92721	6	AX191725	Sequence
c 8	1163	98.5	5612	1	AF043470	Escherich
c 9	328.6	27.8	213494	1	SFPWR100	AL391753 Shigella
c 10	328.6	27.8	221618	1	AF386526	Shigella
c 11	328.6	27.8	221851	1	AF348706	Shigella
c 12	246.6	20.9	3500	1	AF134403	AF134403 Escherich
c 13	159.4	13.5	1642	1	SHFVIRK	D11025 Shigella fl
c 14	159.4	13.5	1642	6	E06635	E06635 Shigella vi
c 15	154.8	13.1	66714	1	AF326777	AF326777 Shigella
c 16	64.2	5.4	6540	1	BACJOJC	L38424 Bacillus su
c 17	64.2	5.4	10801	6	AX100832	Sequence
c 18	64.2	5.4	23775	1	BACYP1A	L47709 Bacillus su
c 19	64.2	5.4	213680	1	BSUB0012	299115 Bacillus su
c 20	59.2	5.0	13334	1	AE003927	AE003927 Xylella f
c 21	57.2	4.8	14243	1	AE007070	AE007070 Mycobacte
c 22	57.2	4.8	23400	1	MTV021	AL021957 Mycobacte
c 23	57	4.8	14734	1	AF498417	AF498417 Pseudomon
c 24	57	4.8	16736	1	AF035937	AF035937 Pseudomon
c 25	54.8	4.6	237221	1	AP003599	AP003599 Nostoc sp
c 26	52.2	4.4	10197	1	AE008327	AE008327 Agrobacte
c 27	52.2	4.4	12006	1	AE009285	AE009285 Agrobacte
c 28	50.6	4.3	20693	1	SMU52844	-U52844 Serratia ma
c 29	49.6	4.2	8594	1	AE013034	AE013034 Thermoana
c 30	49.6	4.2	199450	9	AC005674	AC005674 Homo sapi
c 31	49	4.1	1135	6	AX433280	Sequence
c 32	48.6	4.1	783	6	AX437703	Sequence
c 33	48.2	4.1	253505	1	AP000007	AP000007 Pyrococcu
c 34	48	4.0	35576	1	SC2D46	AL391406 Streptomy
c 35	47.4	4.0	9297	2	AC092541	AC092541 Homo sapi
c 36	47.4	4.0	10991	1	AE010283	AE010283 Pyrococcu
c 37	47.2	4.0	10148	1	AE012967	AE012967 Chloroblu
c 38	46.2	3.9	783	6	AX437696	Sequence
c 39	46.2	3.9	11884	1	AE004745	AE004745 Pseudomon
c 40	46	3.9	11142	1	AE000805	AE000805 Methanoba
c 41	46	3.9	13401	1	AF498404	AF498404 Pseudomon
c 42	45.8	3.9	300600	1	AP005369	AP005369 Thermosyn
c 43	45.4	3.8	10940	1	AE000469	AE000469 Escherich
c 44	45.4	3.8	110000	2	LMFLCHR32_19	Continuation (20 o
c 45	45.4	3.8	176195	1	ECOW89	U00006 E. coli chr
c 46	45.4	3.8	196146	2	AC023609	AC023609 Mus muscu
c 47	45.4	3.8	254993	2	AC078885	AC078885 Mus muscu
c 48	44.8	3.8	125020	9	AF429315	AF429315 Homo sapi
c 49	44.4	3.8	13413	1	AF498405	AF498405 Pseudomon
c 50	44	3.7	12312	1	AE005627	AE005627 Escherich
c 51	44	3.7	318703	1	AP002567	AP002567 Escherich
c 52	43.4	3.7	5280	1	AF105116	AF105116 Streptoco
c 53	43	3.6	10021	1	AE001736	AE001736 Thermotog
c 54	43	3.6	21409	1	AE004175	AE004175 Vibrio ch
c 55	42.8	3.6	36548	1	MSGB1551CS	L78813 Mycobacteri
c 56	42.8	3.6	36548	1	MSGB1554CS	L78814 Mycobacteri
c 57	42.8	3.6	38859	1	MLCB268	AL022602 Mycobacte
c 58	42.8	3.6	348450	1	MLEPRTN4	AL583920 Mycobacte
c 59	42.6	3.6	16349	1	AF498420	AF498420 Pseudomon
c 60	42.6	3.6	299350	1	AP005370	AP005370 Thermosyn
c 61	42.4	3.6	2669	1	AF112861	AF112861 Escherich
c 62	42.4	3.6	20385	1	AE008892	AE008892 Salmonell
c 63	42.2	3.6	264050	1	AL627279	AL627279 Salmonell
c 64	41.6	3.5	36734	1	SC6G10	AL049497 Streptomy
c 65	41.6	3.5	1227	6	AX433356	AX433356 Sequence

66	41.6	3.5	13952	1	AF019375	Escherich	AF019375	Escherich	139	36	3.0	1427	6	AX432342	Sequence	AX432342	Sequence
67	41.4	3.5	17453	1	AF498400	Pseudomon	AF498400	Pseudomon	140	36	3.0	1859	6	AX073342	Sequence	AX073342	Sequence
68	41.4	3.5	341880	1	AP003589	Nostoc sp	AP003589	Nostoc sp	141	36	3.0	13709	1	AE012138	Xanthomon	AE012138	Xanthomon
69	40.8	3.5	9390	6	D87026	Bacillus st	D87026	Bacillus st	142	36	3.0	17971	1	AB033991	Bacillus	AB033991	Bacillus
70	40.6	3.4	14245	6	AX009415	Sequence	AX009415	Sequence	143	36	3.0	157197	2	AC099042	Oryza sat	AC099042	Oryza sat
71	40.6	3.4	129908	6	AX059140	Sequence	AX059140	Sequence	144	36	3.0	170332	2	AC096104	Rattus no	AC096104	Rattus no
72	40.6	3.4	131500	6	RNE603644	Rhizobium	AL603644	Rhizobium	145	36	3.0	345783	1	AP003001	Metorhizo	AP003001	Metorhizo
73	40.4	3.4	10029	1	AE013547	Sequence	AE013547	Sequence	146	35.8	3.0	69806	9	AC026771	Homo sapi	AC026771	Homo sapi
74	40.2	3.4	1894	1	SC0243803	Streptomy	AL133278	Streptomy	147	35.6	3.0	1246	6	AX164174	Sequence	AX164174	Sequence
75	40.2	3.4	34983	1	SCM11	Streptomy	AL133278	Streptomy	148	35.6	3.0	1614	6	AX193704	Streptomy	AX193704	Streptomy
76	40	3.4	24329	1	KPCNPS	Klebsiella	D21242	Klebsiella	149	35.6	3.0	38404	1	SC265	Streptomy	AL035478	Streptomy
77	39.8	3.4	14704	1	AE001970	Deinococc	AE001970	Deinococc	150	35.6	3.0	40337	1	SC5A7	Streptomy	AL031107	Streptomy
78	39.8	3.4	15670	1	AE006951	Mycobacte	AE006951	Mycobacte									
79	39.8	3.4	37218	1	MTIC20G9	Mycobacte	277162	Mycobacteri									
80	39.8	3.4	158781	2	AC024978	Homo sapi	AC024978	Homo sapi									
81	39.8	3.4	162054	2	AC046201	Homo sapi	AC046201	Homo sapi									
82	39.8	3.4	192974	9	AC087565	Homo sapi	AC087565	Homo sapi									
83	39.8	3.4	202726	9	AC092128	Homo sapi	AC092128	Homo sapi									
84	39.6	3.4	2203	8	NCU46088	Neurospora	UA6088	Neurospora									
85	39.6	3.4	2973	8	AF088906	Neurospor	AF088906	Neurospor									
86	39.6	3.4	215889	2	AC092196	Canis fam	AC092196	Canis fam									
87	39.6	3.4	265118	1	CNSPAX06	Pyrococcu	AJ248288	Pyrococcu									
88	39.6	3.4	265118	6	AE041922	Sequence	AE041922	Sequence									
89	39.4	3.3	10206	1	AE013000	Thermoana	AE013000	Thermoana									
90	39.4	3.3	10311	1	AE007621	Clostridi	AE007621	Clostridi									
91	39.4	3.3	10421	1	AE004975	Halobacte	AE004975	Halobacte									
92	38.2	3.3	17013	1	EAAMASL	E.amylolara	X77921	E.amylolara									
93	39.2	3.3	304290	1	CNSPAX05	Pyrococcu	AJ248287	Pyrococcu									
94	39.2	3.3	349980	6	AE041921	Sequence	AE041921	Sequence									
95	39	3.3	11856	1	AE012829	Chlorobiu	AE012829	Chlorobiu									
96	38.8	3.3	10035	1	AE001999	Deinococc	AE001999	Deinococc									
97	38.8	3.3	11172	1	AE010928	Methanosa	AE010928	Methanosa									
98	38.8	3.3	11483	1	AE009930	Pyrocacul	AE009930	Pyrocacul									
99	38.6	3.3	143308	1	D90913	Synechocyst	D90913	Synechocyst									
100	38.4	3.3	68559	2	AC031981	Homo sapi	AC031981	Homo sapi									
101	38.4	3.2	231258	2	AL732467	Mus muscu	AL732467	Mus muscu									
102	38	3.2	34496	1	SCF62	Streptomy	AL121855	Streptomy									
103	37.8	3.2	57818	2	AC017287	Drosophil	AC017287	Drosophil									
104	37.8	3.2	111448	3	AC005334	Drosophil	AC005334	Drosophil									
105	37.8	3.2	166626	3	AC008318	Drosophil	AC008318	Drosophil									
106	37.8	3.2	301639	3	AE003590	Drosophil	AE003590	Drosophil									
107	37.6	3.2	167405	8	AP002483	Oryza sat	AP002483	Oryza sat									
108	37.6	3.2	175565	8	AP003311	Oryza sat	AP003311	Oryza sat									
109	37.4	3.2	33547	8	PCW03G5	Pneumocys	AL592382	Pneumocys									
110	37.2	3.1	1503	8	AF045961	Arabidops	AF045961	Arabidops									
111	37.2	3.1	1533	8	AF045961	Arabidops	AF045961	Arabidops									
112	37.2	3.1	17581	1	AB005901	Streptomy	AB005901	Streptomy									
113	37.2	3.1	125020	9	AF429315	Homo sapi	AF429315	Homo sapi									
114	37.2	3.1	300950	6	AP001516	Bacillus	AP001516	Bacillus									
115	37	3.1	1108	6	AX433315	Sequence	AX433315	Sequence									
116	37	3.1	17576	1	AE004913	Pseudomon	AE004913	Pseudomon									
117	37	3.1	192657	1	AP005377	Thermosyn	AP005377	Thermosyn									
118	37	3.1	199317	2	AC100400	Mus muscu	AC100400	Mus muscu									
119	36.8	3.1	18848	4	CCY14852	Oryctolagus	Y14852	Oryctolagus									
120	36.8	3.1	40352	1	SCC61A	Streptomy	AL356595	Streptomy									
121	36.8	3.1	118507	8	ATF7J8	Arabidops	AL137189	Arabidops									
122	36.8	3.1	300550	1	AP001512	Bacillus	AP001512	Bacillus									
123	36.6	3.1	654	6	AX280161	Sequence	AX280161	Sequence									
124	36.6	3.1	1372	1	AF061562	Mycobacte	AF061562	Mycobacte									
125	36.6	3.1	11119	1	AE013060	Thermoana	AE013060	Thermoana									
126	36.6	3.1	14725	1	AE006956	Mycobacte	AE006956	Mycobacte									
127	36.6	3.1	40051	1	MSG224	AD000004	AD000004	AD000004									
128	36.6	3.1	40838	1	MTY25D10	Mycobacteri	Z95558	Mycobacteri									
129	36.4	3.1	175127	2	AC121333	Homo sapi	AC121333	Homo sapi									
130	36.4	3.1	183914	2	AC069331	Homo sapi	AC069331	Homo sapi									
131	36.2	3.1	741	6	AX280220	Sequence	AX280220	Sequence									
132	36.2	3.1	1704	6	AX193681	Sequence	AX193681	Sequence									
133	36.2	3.1	10732	6	E32986	Gene encodi	E32986	Gene encodi									
134	36.2	3.1	67521	2	AC103845	Homo sapi	AC103845	Homo sapi									
135	36.2	3.1	150052	9	AC018525	Homo sapi	AC018525	Homo sapi									
136	36.2	3.1	178736	9	AC103853	Homo sapi	AC103853	Homo sapi									
137	36.2	3.1	181688	2	AP005354	Rattus no	AP005354	Rattus no									
138	36.2	3.1	192581	2	AC121044	Sequence	AC121044	Sequence									

ALIGNMENTS

RESULT 1	AX011298	Sequence 2 from Patent WO9955908.	1181 bp	DNA	Linear	PAT 06-SEP-2000
LOCUS	AX011298	Sequence 2 from Patent WO9955908.	1181 bp	DNA	Linear	PAT 06-SEP-2000
DEFINITION	AX011298	Sequence 2 from Patent WO9955908.	1181 bp	DNA	Linear	PAT 06-SEP-2000
ACCESSION	AX011298	Sequence 2 from Patent WO9955908.	1181 bp	DNA	Linear	PAT 06-SEP-2000
VERSION	AX011298.1	GI:9997848				
KEYWORDS		Escherichia coli.				
SOURCE		Escherichia coli.				
ORGANISM		Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
REFERENCE		1 (bases 1 to 1181)				
AUTHORS		Thierry, D., Frechon, D.T. and Laure, F.C.				
TITLE		Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)				
JOURNAL		Patent: WO 9955908-A 2 04-NOV-1999;				
FEATURES		Thierry DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)				
Source		Location/Qualifiers				
		i. 1181				
		/organism="Escherichia coli"				
		/db_xref="taxon:562"				
BASE COUNT	305 a	317 c	277 g	282 t		
ORIGIN						
Query Match	100.0%;	Score 1181;	DB 6;	Length 1181;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1181;	Conservative	0;	Mismatches	0;	Gaps	0;
QY	1	CTGCAGGAGATGCAAAAAGCAAAAATTAATAATTCCTCCATCCCGCGCTCCAGCTG	60			
Db	1	CTGCAGGAGATGCAAAAAGCAAAAATTAATAATTCCTCCATCCCGCGCTCCAGCTG	60			
QY	61	AAAGTAGGCGCTTCTGTCGGGTATTTAAATGCAATGACCGTCCCGCTATTTAAACAATG	120			
Db	61	AAAGTAGGCGCTTCTGTCGGGTATTTAAATGCAATGACCGTCCCGCTATTTAAACAATG	120			
QY	121	TGATAAATACCTCCGTACCGGAAACCGCTGACCAAAATTCGGGCTGAAAGAGGATCC	180			
Db	121	TGATAAATACCTCCGTACCGGAAACCGCTGACCAAAATTCGGGCTGAAAGAGGATCC	180			
QY	181	GGCGTTATCTGTGTCATTTCCCTTAGCTGACTAGCCAGACAGACAATGATCTGTGCGG	240			
Db	181	GGCGTTATCTGTGTCATTTCCCTTAGCTGACTAGCCAGACAGACAATGATCTGTGCGG	240			
QY	241	TTCCTTTAATCAAAACCGGTACTCAATATCTTCTCTGCGCTGGCTGCCATCATCCGGA	300			
Db	241	TTCCTTTAATCAAAACCGGTACTCAATATCTTCTCTGCGCTGGCTGCCATCATCCGGA	300			
QY	301	ACGCTTCGGTGGGATAAAAATCGGCAGTGGCGGTCCATGCAGACATCCGCCA	360			
Db	301	ACGCTTCGGTGGGATAAAAATCGGCAGTGGCGGTCCATGCAGACATCCGCCA	360			
QY	361	CGGGTAACAGCGGTCCTCTGACATTTCTTCTGAATGACATCAGGATCCCGCCGCTTCAC	420			
Db	361	CGGGTAACAGCGGTCCTCTGACATTTCTTCTGAATGACATCAGGATCCCGCCGCTTCAC	420			

QY 421 TGCGGATACGGGACCGCGGAGACTGACGCTTCACCGAGTACCATACCAACGCTCAT 480
 Db 421 TGCGGATACGGGACCGCGGAGACTGACGCTTCACCGAGTACCATACCAACGCTCAT 480
 QY 481 TTTCGGAAGCATGACACACACTCGGCAATCCGGTAGACCGGTACGCTGGGAAAGG 540
 Db 481 TTTCGGAAGCATGACACACACTCGGCAATCCGGTAGACCGGTACGCTGGGAAAGG 540
 QY 541 CACCTGCCATTAAACACATCCGCTCATCCAGGTGTTCTGCTGACGAGCAGGTG 600
 Db 541 CACCTGCCATTAAACACATCCGCTCATCCAGGTGTTCTGCTGACGAGCAGGTG 600
 QY 601 CTTGCTATTCTACCGCCCGCGGCCCCACACAGCAGCAGCAATGATTTCCCTTCATCT 660
 Db 601 CTTGCTATTCTACCGCCCGCGGCCCCACACAGCAGCAGCAATGATTTCCCTTCATCT 660
 QY 661 TCAGCTGATACATACACGACGACATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCA 720
 Db 661 TCAGCTGATACATACACGACGACATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCA 720
 QY 721 CTTGAACGATAGCGGAACATTGCTGCTGATGACGACGCGGTGATGACGAGGTA 780
 Db 721 CTTGAACGATAGCGGAACATTGCTGCTGATGACGACGCGGTGATGACGAGGTA 780
 QY 781 ACGGTCGCATGCTTCATTATGCAATGCGGGCCAGTCGAACCCGGTGAATAACCGTTA 840
 Db 781 ACGGTCGCATGCTTCATTATGCAATGCGGGCCAGTCGAACCCGGTGAATAACCGTTA 840
 QY 841 CCGGTCCTGACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGCACACAA 900
 Db 841 CCGGTCCTGACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGCACACAA 900
 QY 901 TGAATCACACAGATAATTGAGGAAACGTTCTGCTCTACGCGGTGATGAGTGTGTTT 960
 Db 901 TGAATCACACAGATAATTGAGGAAACGTTCTGCTCTACGCGGTGATGAGTGTGTTT 960
 QY 961 GTCTGACATAGTGAACGGTGACACATATCAGACGCTCAGTCTGCTATATTAATGTT 1020
 Db 961 GTCTGACATAGTGAACGGTGACACATATCAGACGCTCAGTCTGCTATATTAATGTT 1020
 QY 1021 CATGGCCACTATGGCAGATGACACATCAGGTTTAAATCCCGATAATCCCGTGAAGTC 1080
 Db 1021 CATGGCCACTATGGCAGATGACACATCAGGTTTAAATCCCGATAATCCCGTGAAGTC 1080
 QY 1081 TGAGGATGAAGGAGGTGAAGCTGTTCTGAAAGGAATAAAGTGACATCATGCCCTC 1140
 Db 1081 TGAGGATGAAGGAGGTGAAGCTGTTCTGAAAGGAATAAAGTGACATCATGCCCTC 1140
 QY 1141 TTTTCTGCTCCGAGCAATTTTACTTTTTTCTCTGCAG 1181
 Db 1141 TTTTCTGCTCCGAGCAATTTTACTTTTTTCTCTGCAG 1181

RESULT 2
 EC74P0157/c
 LOCUS EC74P0157 7395 bp DNA linear BCT 07-JAN-1999
 DEFINITION E.coli 7.4 kb DNA from plasmid p0157.
 ACCESSION Y11275
 VERSION Y11275.1 GI:4127812
 KEYWORDS msbB gene; unidentified protein.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 7395)
 Brander, W.
 Unpublished
 2 (bases 1 to 7395)
 Brander, W.
 Direct Submission
 Submitted (14-FEB-1997) W. Brander, Universitaet Wuerzburg,
 Institut fuer Hygiene und Mikrobiologie, Josef-Schneider-Strasse 2,

D- 97080 Wuerzburg, FRG
 Related sequences X97542, M77039, X02311.
 Location/Qualifiers
 1. .7395
 /organism="Escherichia coli"
 /strain="EDL933"
 /serotype="O157:H7"
 /db_xref="taxon:562"
 /clone="psm7"
 /plasmid="p0157"
 774. .1595
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="CAA72138.1"
 /db_xref="GI:4127813"
 /translation="MLNARHLPLVLMYHVSRCPLVTLSPETFEOMKVLADNNWRTV
 TSELEVFYGGTLPKRSYMLTFDGYLDNWLRAWPVLAKEYNLHAHFLITGLIGGN
 VSRQDEYSHRDCERLIAADNRSEVLRWSEMRDSSLVEFHVHTSHKRWDRLS
 VRAQECRLMKEDILGKQCLEKLGFCSCWPEGYNRYDINAGLKGFSYLTT
 ERMNCPENGSLRIGRISTKREHSGWLKRLFYTTPLFSSVLALHKGRLPDN"
 1595. .2701
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="CAA72139.1"
 /db_xref="GI:4127814"
 /translation="MMKILFTSSSDIGGOELQAQMTALQKOGHSLVLAEREKSKI
 APEARKGHDTVFPFRNSLHPLRLRIIGFEPDLVICHSGHDSNLAGLSRLIC
 CHFSIVRQKTYITRTYFSLNYLDFIVPSSAMMAHLMAGVTPVTVIIPGDDW
 PALHNEAMPLPHIHAASADNVLIVQVGLMRPEKHEFMRVLYOUMEGKSPR
 WLIVGAREYEARLQOQTEHGMGDLVLMAGALFALPVRIASVYVAMPSENEAFGM
 VLAESVGPVPIASITGGTIPVIOKNTGCTLLPVGDSVAWTGALRDFLSRPERFRM
 AASARDIEYRFDINRTAQIIVSLASQAKGKCNR"
 2791. .4512
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="CAA72140.1"
 /db_xref="GI:4127815"
 /translation="MHLNTGONRPTFSWALGWAIFYFGFFSTLLOVIFSSGSGTN
 GIRDLSLCLMLIPVLYPDRKIIIAAVVGLLWGTSLAACLYLYLGHFSQVLF
 VMFETNARAGEYFSQYFSLKLLLSLYTAVSVFLWTLRLPVYILPWHRIYSVLL
 YALLHPVLKLSIROEPLNDLGLKASREPAAPQFVSSYYOHOQNALTFLLNE
 NSALPLGLNRDESERPTLVIGESTQREMSLYGLYRETTPELDALRKTDGLT
 VNNVVASRPTTEALQALTFANEKNPDLYTPQLMMMKQAGKYKWTNQQTIT
 ARNTMLTVESSQTDROYVMNQRTQSAVEDYVNLKPFREVLNDPAKKLIIVHLLGT
 HKYKYRYPGOGGREGITGHIPTGLNAKELEVYNDYDNANVFNHVASLINDFRAT
 APDGLLYFSDHGEVEYDTPPYKTQGRNEDNPRPMYTPVFLWTSEKHAHAPRDS
 QYVDRKISLAELIHTWSDLAGLTYDGYDPTSLVNPQFRETTRWIGNPKYKKNGLTDFD
 TLPYGEF"
 4586. .5584
 /gene="msbB"
 4586. .5584
 /gene="msbB"
 /codon_start=1
 /transl_table=11
 /product="msbB protein"
 /protein_id="CAA72141.1"
 /db_xref="GI:4127816"
 /translation="MVPPSAVLCYHNEISRQIPVNMKNITETIIPRNLTLCPRTWM
 TWIGIGICVFAMVPPALRDLPLGLKGLMVLGRKLSARALINLSLCPPEYSDKEKE
 NIVDAMFATASMAVLMALALSGDPKISHRIWNGLEIVKMAQNKQKVIYFLVPHAW
 GVDIIPAMLAASGRKMAAFHMQRPVYVWNSVRRFGKLRANDGJASVRSVR
 QGYWGYLPDQDHGPEFSFADPFATYKATLPVIGRLSRISGARIIPLPVYDGKTHH
 LTHVSPPLAIRQKSDAHARQINEVVEVFRPHEQYTWILKLLATKREGEDPY"
 5645. .5820
 /note="IS3-like sequence"
 6465. .57395
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="CAA72142.1"
 misc_feature
 CDS

```
/db_xref="GI:4127817"
/translation="MNTKMERWRTPMKLYLSCTLAPLAIGVFSATAADNNSAIYF
NTSOPINDLQCSLAIEVFAQSILPAHPKREGDSQPHLTSRLKSLILVFPVXRADRTF
VOVARDNNKILGTLTYLPPSLPTDIYHLGVPEGGIDFPHNGTKKIINTVAEVN
KLSDAGSSIHSHLTHNALVEIHTANGRWVDIYLPQDPLECKMVFVSSAGYSTV
FYGRKVTLSVGNLTFLPKYNGWFRSGELENNRIYTAQHINSAELPAHWIVPLNLV
IKGNLSGRUNDIKIGAPGBELLHTIDIGMLTTP"
BASE COUNT 1929 a 1575 c 1815 g 2076 t
ORIGIN

Query Match 98.7%; Score 1166.2; DB 1; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
DB 2875 TGGAAAAAGCCAAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2816
QY 71 TGTCTGTCCGGTATTAAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 TGTCTGTCCGGTATTAAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 131 CTCGGTACCGGAAACCGCTGAACAAATTCGGCTGAAAGAGGATCCGCGTATCT 190
DB 2755 CTCGGTACCGGAAACCGCTGAACAAATTCGGCTGAAAGAGGATCCGCGTATCT 2696
QY 191 GTTTCCTGCGGTATTAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 GTTTCCTGCGGTATTAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 131 CTCGGTACCGGAAACCGCTGAACAAATTCGGCTGAAAGAGGATCCGCGTATCT 190
DB 2755 CTCGGTACCGGAAACCGCTGAACAAATTCGGCTGAAAGAGGATCCGCGTATCT 2696
QY 191 GTTTCCTGCGGTATTAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 GTTTCCTGCGGTATTAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 251 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCGGAAGCGTTCGG 310
DB 2635 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCGGAAGCGTTCGG 2576
QY 311 TCGGGATAAAAAATCGCGAGTCGCGGTCATGACAGACACATCCCGGAGCGTTCGG 370
DB 2575 TCGGGATAAAAAATCGCGAGTCGCGGTCATGACAGACACATCCCGGAGCGTTCGG 2516
QY 371 CGTCCCTGCATCTTCTGAATGACATGACAGGATCCGCGCTCTCACTGGCGATAAC 430
DB 2515 CGTCCCTGCATCTTCTGAATGACATGACAGGATCCGCGCTCTCACTGGCGATAAC 2456
QY 431 GGGCAGCGCGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 490
DB 2455 GGGCAGCGCGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 2396
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAGCGGTAGCGGTAGCGGTAGCGGT 550
DB 2395 CATGACCACCACTGGCAATCCGGTAGACCGGTAGCGGTAGCGGTAGCGGTAGCGGT 2336
QY 551 TAACACATCTCCGTCATCTCCAGGTGTTCTGTCTGTGACGACAGAGTGTCTGTATTC 610
DB 2335 TAACACATCTCCGTCATCTCCAGGTGTTCTGTCTGTGACGACAGAGTGTCTGTATTC 2276
QY 611 TTCAGCGCGCGCGCCACGACGCGGAGGAAATGATTCCTTCATCTTCAGTCTATA 670
DB 2275 TTCAGCGCGCGCGCCACGACGCGGAGGAAATGATTCCTTCATCTTCAGTCTATA 2216
QY 671 CAATACAGCGAGCAATAATTCATCTCTTTTCGGGACGTAGCATCCCGACCTGAACGAT 730
DB 2215 CAATACAGCGAGCAATAATTCATCTCTTTTCGGGACGTAGCATCCCGACCTGAACGAT 2156
QY 731 AAGCGGAACATGTCTGTATGACGCGCGGAGGATGACAGGAGTAAAGGTCGCAT 790
DB 2155 AAGCGGAACATGTCTGTATGACGCGCGGAGGATGACAGGAGTAAAGGTCGCAT 2096
QY 791 GCCTTCATTTGCAATCGCGGCGAGTCGAACCGCGGTGGAATACCGTTCACCGTTCCT 850
DB 2095 GCCTTCATTTGCAATCGCGGCGAGTCGAACCGCGGTGGAATACCGTTCACCGTTCCT 2036
QY 851 GACACCTTCGCGCATGACATGCGGCATCATGGGTGAGATAGGACACAAATGAATCACA 910
DB 2035 GACACCTTCGCGCATGACATGCGGCATCATGGGTGAGATAGGACACAAATGAATCACA 1976
```

```
911 CAGATAAATTCAGGAAAAAGCTTCTGGTCTTACGGGTGATGTAGGTTTTTTGTCTGACAAT 970
1975 CAGATAAATTCAGGAAAAAGCTTCTGGTCTTACGGGTGATGTAGGTTTTTTGTCTGACAAT 1916
971 AGTGAACCGGTGACAGCATATCAGACGGCTCAGTCTCTGCTATATTACTGTCTAGTCCACT 1030
1915 ACTGAACCGGTGACAGCATATCAGACGGCTCAGTCTCTGCTATATTACTGTCTAGTCCACT 1856
1031 ATGGCAGATGACAGCATCAGGTTTAAATTCGCCGATAATCCCGTGAAGTCTCAGGATGA 1090
1855 ATGGCAGATGACAGCATCAGGTTTAAATTCGCCGATAATCCCGTGAAGTCTCAGGATGA 1796
1091 AGAAGGTGAAGGCTTCTCTGAAAGGAATAAAGTGAACATCATGCCCTCTTTTCTGGC 1150
1795 AGAAGGTGAAGGCTTCTCTGAAAGGAATAAAGTGAACATCATGCCCTCTTTTCTGGC 1736
1151 TTCCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
1735 TTCCCGGAGCAATTTTACTTTTTTCTCTGCAG 1705

RESULT 3
AX191728/c
LOCUS AX191728 7395 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 10 from Patent WO0149775.
ACCESSION AX191728
VERSION AX191728.1 GI:15209897
KEYWORDS Escherichia coli.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
REFERENCE 1 (bases 1 to 7395)
AUTHORS Iversen,P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 10 12-JUL-2001;
AVI Biopharma, Inc. (US)
FEATURES
source
Location/Qualifiers
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 1929 a 1575 c 1815 g 2076 t
ORIGIN

Query Match 98.7%; Score 1166.2; DB 6; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
DB 2875 TGGAAAAAGCCAAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2816
QY 71 TGTCTGTCCGGTATTAAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 TGTCTGTCCGGTATTAAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 131 CTCGGTACCGGAAACCGCTGAACAAATTCGGCTGAAAGAGGATCCGCGTATCT 190
DB 2755 CTCGGTACCGGAAACCGCTGAACAAATTCGGCTGAAAGAGGATCCGCGTATCT 2696
QY 191 GTTTCCTGCGGTATTAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 GTTTCCTGCGGTATTAATGACATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
QY 251 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCGGAAGCGTTCGG 310
DB 2635 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCGGAAGCGTTCGG 2576
QY 311 TCGGGATAAAAAATCGCGAGTCGCGGTCATGACAGACACATCCCGGAGCGTTCGG 370
DB 2575 TCGGGATAAAAAATCGCGAGTCGCGGTCATGACAGACACATCCCGGAGCGTTCGG 2516
QY 371 CGTCCCTGCATCTTCTGAATGACATGACAGGATCCGCGCTCTCACTGGCGATAAC 430
DB 2515 CGTCCCTGCATCTTCTGAATGACATGACAGGATCCGCGCTCTCACTGGCGATAAC 2456
QY 431 GGGCAGCGCGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 490
DB 2455 GGGCAGCGCGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTCCGAGG 2396
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAGCGGTAGCGGTAGCGGTAGCGGT 550
DB 2395 CATGACCACCACTGGCAATCCGGTAGACCGGTAGCGGTAGCGGTAGCGGTAGCGGT 2336
QY 551 TAACACATCTCCGTCATCTCCAGGTGTTCTGTCTGTGACGACAGAGTGTCTGTATTC 610
DB 2335 TAACACATCTCCGTCATCTCCAGGTGTTCTGTCTGTGACGACAGAGTGTCTGTATTC 2276
QY 611 TTCAGCGCGCGCGCCACGACGCGGAGGAAATGATTCCTTCATCTTCAGTCTATA 670
DB 2275 TTCAGCGCGCGCGCCACGACGCGGAGGAAATGATTCCTTCATCTTCAGTCTATA 2216
QY 671 CAATACAGCGAGCAATAATTCATCTCTTTTCGGGACGTAGCATCCCGACCTGAACGAT 730
DB 2215 CAATACAGCGAGCAATAATTCATCTCTTTTCGGGACGTAGCATCCCGACCTGAACGAT 2156
QY 731 AAGCGGAACATGTCTGTATGACGCGCGGAGGATGACAGGAGTAAAGGTCGCAT 790
DB 2155 AAGCGGAACATGTCTGTATGACGCGCGGAGGATGACAGGAGTAAAGGTCGCAT 2096
QY 791 GCCTTCATTTGCAATCGCGGCGAGTCGAACCGCGGTGGAATACCGTTCACCGTTCCT 850
DB 2095 GCCTTCATTTGCAATCGCGGCGAGTCGAACCGCGGTGGAATACCGTTCACCGTTCCT 2036
QY 851 GACACCTTCGCGCATGACATGCGGCATCATGGGTGAGATAGGACACAAATGAATCACA 910
DB 2035 GACACCTTCGCGCATGACATGCGGCATCATGGGTGAGATAGGACACAAATGAATCACA 1976
```

```
QY 371 CGTCCCTGTACATTTCTTCTGATGACATCAGGATCCGCCGCTCTCACTGGCGATAAC 430
Db 2515 CGTCCCTGTACATTTCTTCTGATGACATCAGGATCCGCCGCTCTCACTGGCGATAAC 2456
QY 431 GGCACGCCGGAGACTGACGCTTCACCCAGTACCATACCAAGCGCTTCATTTTCCGAAG 490
Db 2455 GGCACGCCGGAGACTGACGCTTCACCCAGTACCATACCAAGCGCTTCATTTTCCGAAG 2396
QY 491 CATGACCCACACACTGGCAATCCGGTAGCCGTAACGCTGGGAAAGGCGACCTGCCAT 550
Db 2395 CATGACCCACACACTGGCAATCCGGTAGCCGTAACGCTGGGAAAGGCGACCTGCCAT 2336
QY 551 TTAACATCTCCGCTCATTTCCAGGTGTTCTGTCTGCTACGACGAGCTGCTTCGTATTC 610
Db 2335 TTAACATCTCCGCTCATTTCCAGGTGTTCTGTCTGCTACGACGAGCTGCTTCGTATTC 2276
QY 611 TTCACGCCCGGGCCGCCACACGAGCGGAAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 2275 TTCACGCCCGGGCCGCCACACGAGCGGAAATGATTTCCCTTCCATCTTCAGCTGATA 2216
QY 671 CAATACACGACGATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
Db 2215 CAATACACGACGATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 2156
QY 731 AAGCGAACAATGTCTGTGATGACGCCGAGCGGTGGATATGAGGGGTAAAGGTGCGAT 790
Db 2155 AAGCGAACAATGTCTGTGATGACGCCGAGCGGTGGATATGAGGGGTAAAGGTGCGAT 2096
QY 791 GGCTTCATTTATGCAATCGCGGCCAGTCGAAACCCGGTGAATTAACCGTTACCGGTCTCT 850
Db 2095 GGCTTCATTTATGCAATCGCGGCCAGTCGAAACCCGGTGAATTAACCGTTACCGGTCTCT 2036
QY 851 GACACCTTCGCCATCAGATGCGCCATCATGGTGTAGATAGGCACACAATGAATACACA 910
Db 2035 GACACCTTCGCCATCAGATGCGCCATCATGGTGTAGATAGGCACACAATGAATACACA 1976
QY 911 CAGATAATTCAGGAAAGCTTCTGTCTTACGGGTGATGATAGTTTCTGTGACAT 970
Db 1975 CAGATAATTCAGGAAAGCTTCTGTCTTACGGGTGATGATAGTTTCTGTGACAT 1916
QY 971 AGTGAAGCGGTGACAGCATATCAGCGCTCAGTCTCTGTATATTACTGTCTATGCGCACT 1030
Db 1915 AGTGAAGCGGTGACAGCATATCAGCGCTCAGTCTCTGTATATTACTGTCTATGCGCACT 1856
QY 1031 ATGGCAGATGACAGATCAGTTTAAATTCGCCGATAATCCGTCGAAGTCTGAGGTGGA 1090
Db 1855 ATGGCAGATGACAGATCAGTTTAAATTCGCCGATAATCCGTCGAAGTCTGAGGTGGA 1796
QY 1091 AGGAAGGTGAAGCTGTTCTGTAAGGATAAAGTGACATCATGCCCCCTCTTTTCTGCG 1150
Db 1795 AGGAAGGTGAAGCTGTTCTGTAAGGATAAAGTGACATCATGCCCCCTCTTTTCTGCG 1736
QY 1151 TTCGGGAGCAATTTTACTTTTCTCTGCGAG 1181
Db 1735 TTCGGGAGCAATTTTACTTTTCTCTGCGAG 1705
```

RESULT 4

```
AF074613/c 92077 bp DNA circular BCT 04-NOV-1998
LOCUS
DEFINITION Escherichia coli O157:H7 plasmid pO157, complete sequence.
ACCESSION AF074613
VERSION AF074613.1 GI:3822114
KEYWORDS
SOURCE Escherichia coli O157:H7.
ORGANISM Escherichia coli O157:H7
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
REFERENCE 1 (bases 1 to 92077)
AUTHORS Burland,V., Shao,Y., Perna,N.T., Plunkett,G., Sofia,H.J. and
Blattner,F.R.
TITLE The complete DNA sequence and analysis of the large virulence
plasmid of Escherichia coli O157:H7
```

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Nucleic Acids Res. 26 (18), 4196-4204 (1998)
98391744
9722640
2 (bases 1 to 92077)
Burland,V., Shao,Y., Perna,N.T., Plunkett,G. III, Sofia,H.J. and
Blattner,F.R.
Direct Submission
Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. 92077
/organism="Escherichia coli O157:H7"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:83334"
/lab_host="Escherichia coli C600"
/plasmid="pO157"
1. 561
/gene="fno"
/note="L7001"
1. 561
/gene="fno"
/note="97 pct identical amino acid sequence and equal
length to F101_ECOLI SW: P22707"
/codon_start=1
/transl_table=11
/product="fertility inhibition protein (conjugal transfer
repressor)"
/protein_id="AAC70069.1"
/db_xref="GI:3822115"
/translation="MAEQKRPVLTLRKTEGETPVRSRKTIINVTTPPKVKVKQKLA
EKAAREALEAKQAQALSYLNLTQDEAVNTLKPWPGLDGDTPLRLACGIRD
VLEDAQNIPLSHKRLRAKAITRESYLCAMKAGACRYDTGYVYTHISQEEBA
YAAARLDKIRQRNRIKAEQLQAVLDER"
653. 681
/note="predicted sigma 70 promoter; score of 52%"
/gene="L7002"
698. 949
/gene="L7002"
698. 949
/gene="L7002"
/note="54 pct identical (1 gap) to 23 residues of an
approx. 269 aa protein EXOK_RHIME SW: P33693"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC70070.1"
/db_xref="GI:3822116"
/translation="MDSEYVHGTVRSVTSVPEAGPLFWKSVSDAGWKRQKHGDLPLV
HGLTGSUPEKGLNTATGAAEEGNEKSSLSHYRDSG"
1025. 1052
/note="predicted sigma 70 promoter; score of 62%"
1151. 1612
/gene="L7003"
1151. 1612
/gene="L7003"
/note="98 pct identical and equal length to YFI3_ECOLI SW:
Q99342"
/codon_start=1
/transl_table=11
/product="hypothetical protein 15.6 kDa protein in fno 3'
region precursor"
/protein_id="AAC70071.1"
/db_xref="GI:3822117"
/translation="MRKYIPLVLFISWPVLCADIHGRVVRVLDGDTIEVMSRKAVR
IRLVNIDAPEKKQDYGWSTDMKSLVACKTVITYTFQDRDRGRMLGQVYAPGMVNV
QFMVRAGAAWYEQYINTDPVLPVQLNEARQQKRLGWSADPVPFWIWRKR"
1390. 1419
/note="predicted sigma 70 promoter; score of 56%"
1658. 1867
/gene="L7004"
1658. 1867
/gene="L7004"
/note="55 pct identical (0 gaps) to 66 residues of an

promoter

gene

CDS

promoter

gene

CDS

promoter

gene

CDS

```

approx. 72 aa protein HHA_ECOLI SW: P23870"
/codon_start=1
/transl_table=11
/product="putative hemolysin expression modulating
protein"
/protein_id="AAC70072.1"
/db_xref="GI:3822118"
/translation="MEXTQEWLYQLRCCSSVNTLEKIIHKNRDSLSTSERESFNAA
DHRAELITGKLYDRIPKEIKVYR"
1899..1927
/feature="predicted sigma 70 promoter; score of 56%"
1905..2243
/feature="L7005"
1905..2243
/feature="L7005"
/feature="98 pct identical (0 gaps) to 57 residues of an
approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF
4, TRSW: Q52340"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC70073.1"
/db_xref="GI:3822119"
/translation="MKLIIFILVLIIAALLIRILSVNOHSPLLMOLHAAGIRGTD
AERTLSGSEVDASHLERRRRDILLEVLQRGIPLLRSHDARKLLQMTGEWLNTTG
AAQSPHRS"
complement(240..2327)
/feature="initiation site; Escherichia coli plasmid R100 ss1B
gene; 99 pct identical to ss1B locus R10SS1B accession
D90185"
2416..2421
/feature="100 pct identical (0 gaps) to the -35 region at
23..28 locus ECNR1REP accession X02302"
2439..2445
/feature="100 pct identical (0 gaps) to the -10 region at
46..52 locus ECNR1REP accession X02302"
2473..2478
/feature="100 pct identical (0 gaps) to RBS at 80..85 locus
ECNR1REP accession X02302"
2483..2737
/feature="cpb2"
2483..2737
/feature="L7006"
2483..2737
/feature="cpb2"
/feature="98 pct identical amino acid sequence and equal
length to CPB2_ECOLI SW: P03847"
/codon_start=1
/transl_table=11
/product="CcpB protein (RepA2 protein)"
/protein_id="AAC70074.1"
/db_xref="GI:3822120"
/translation="MSQTENAVTSSSGAKRAYRKGPNPLSDAEKQRLSVARKRAFKEV
KVLEPKYKALMQMCHEDGLTQAEVLTLIKSEAQRKV"
2754..2782
/feature="predicted sigma 70 promoter; score of 59%"
2973..3047
/feature="L7007"
2973..3047
/feature="L7007"
/feature="91 pct identical (0 gaps) to 24 residues of an
approx. 24 aa protein uORF P30REPFIC, accession M16167,
translationally coupled to replication initiation protein"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70075.1"
/db_xref="GI:3822121"
/translation="MLGKVDQDFLCSILLRIYVAGWCD"
3032..3035
/feature="100 pct identical (0 gaps) to RBS at 641..644 locus
ECNR1REP accession X02302"
3040..3897
/feature="rep2"

```

```

/feature="L7008"
3040..3897
/feature="rep2"
/feature="96 pct identical amino acid sequence and equal
length to REP2_ECOLI SW: P03066"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/translation="MTDIQOTVYROVKPNPNVETPRGAGTLKCEKLMKAVGFTSR
FDFAIHVAHRSKGLRRMPVLRRAIDALLOGLCFHYDPLANRVQCSITFLAIECG
LATESAGKLSITRATLTFUSELGLITYQTEYDPLIGCIPTDITTFALFALDV
SEDAVAARRSRVWENLRKKQGLDTGMDELIKAWFRVFRFRSTQELKSRGKIK
RARARDAGRRERQDIVTLVKQLTREISGREFRANREAVKREVERVERMILSRNRN
YSRLATASP"
3847..4095
/standard_name="CIS"
/feature="88 pct identical to locus ECIS accession X12587,
required for cis-activation of orf1 by the replication
initiation protein"
4072..4080
/feature="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1REP at (1882..1690) accession X02302"
4084..4232
/standard_name="oriR"
/feature="89 pct identical to oriR (1094..1242); minimum
segment for replication of E. coli IncFII plasmid NR1
ECREPAL X12776"
/direction=right
4169..4197
/feature="predicted sigma 70 promoter; score of 56%"
4259..4453
/feature="L7009"
4259..4453
/feature="L7009"
/feature="95 pct identical to (0 gaps) 64 residues of a 128
aa protein REP44 locus ECRC1 accession V00351"
/codon_start=1
/transl_table=11

```

```

Query Match      98.7%; Score 1166.2; DB 1; Length 92077;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGTGAAGTAGGCC 70
DB 19424 TGGAAAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGTGAAGTAGGCC 19365
QY 71 TCTTCTGTCGGTATTAAATGCATGACCGTCCCGCTATTAAACAATGTGATAAATTA 130
DB 19364 TCTTCTGTCGGTATTAAATGCATGACCGTCCCGCTATTAAACAATGTGATAAATTA 19305
QY 131 CTCGGTTACCGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCCGGTATCT 190
DB 19304 CTCGGTTACCGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCCGGTATCT 19245
QY 191 GTTGCAATTTCCCTTAGCTAGTACGACAGACACAATGATCTGTGCGCTTCTGTTAAT 250
DB 19244 GTTGCAATTTCCCTTAGCTAGTACGACAGACACAATGATCTGTGCGCTTCTGTTAAT 19185
QY 251 ATCAACACGGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCTCCGGAAGCGTCCGG 310
DB 19184 ATCAACACGGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCTCCGGAAGCGTCCGG 19125
QY 311 TCGGGATAAAAAATCGCCAGTGGCCGGTCCATGACACACATCCCCCAGGGGTAAACAG 370
DB 19124 TCGGGATAAAAAATCGCCAGTGGCCGGTCCATGACACACATCCCCCAGGGGTAAACAG 19065
QY 371 CGTCCCTGTACATTTCTTCTGAATGACATCAGGGATCCCGCGCTCTCACTGGCGATAAC 430
DB 19064 CGTCCCTGTACATTTCTTCTGAATGACATCAGGGATCCCGCGCTCTCACTGGCGATAAC 19005

```



```
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTTCGAAGG 490
Db 19004 GGGCAGCCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTTCGAAGG 18945
QY 491 CATGACACACACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 550
Db 18944 CATGACACACACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 18885
QY 551 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTTCGTAATTC 610
Db 18884 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTTCGTAATTC 18825
QY 611 TTACAGCCGGGCCCCACACGAGCGAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 18824 TTACAGCCGGGCCCCACACGAGCGAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 18765
QY 671 CAATACACGAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
Db 18764 CAATACACGAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 18705
QY 731 AAGCGGAACATCTGCTGTAGTACGAGCCAGCGGTGGATATGAGGGGTAAACGGTGCAT 790
Db 18704 AAGCGGAACATCTGCTGTAGTACGAGCCAGCGGTGGATATGAGGGGTAAACGGTGCAT 18645
QY 791 GGCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCTCT 850
Db 18644 GGCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCTCT 18585
QY 851 GACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGATAGGACCAACATGAATACACA 910
Db 18584 GACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGATAGGACCAACATGAATACACA 18525
QY 911 CAGATAATTCAGGAAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 970
Db 18524 CAGATAATTCAGGAAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 18465
QY 971 AGTGAAGCGGTGACAGCATATCAGCGCTGCTGCTGCTATATTTACTGTCATGGCCACT 1030
Db 18464 ACTGAAGCGGTGACAGCATATCAGCGCTGCTGCTGCTATATTTACTGTCATGGCCACT 18405
QY 1031 ATGGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCGTCGAGTCTGAGGATGA 1090
Db 18404 ATGGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCGTCGAGTCTGAGGATGA 18345
QY 1091 AGGAAGTGAAGCTGTTCTGAAAGGAATAAAGTGACATCATGCCCCCTTTTCTGGC 1150
Db 18344 AGGAAGTGAAGCTGTTCTGAAAGGAATAAAGTGACATCATGCCCCCTTTTCTGGC 18285
QY 1151 TTCCGGAGCAATTTTACTTTTCTCTGCAG 1181
Db 18284 TTCCGGAGCAATTTTACTTTTCTCTGCAG 18254
```

```
RESULT 5
AX191727/c 92077 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 9 from Patent WO0149775.
ACCESSION AX191727
VERSION AX191727.1 GI:15209896
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
AUTHORS Iversen,P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 9 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source
1. .92077
/organism="Escherichia coli"
/db_xref="taxon:562"
```

```
BASE COUNT 24693 a 20133 c 23728 g 23523 t
ORIGIN
Query Match 98.7%; Score 1166.2; DB 6; Length 92077;
Best Local Similarity 99.7%; Pred No. 0;
Matches 1188; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAAATAAAAAATGCCCCATCCAGCGCTCCAGCTGAAAGTAGGCC 70
Db 19424 TGGAAAAAGCCAAAATAAAAAATGCCCCATCCAGCGCTCCAGCTGAAAGTAGGCC 19365
QY 71 TGTCTCTGTCGGGTATTAAATGCAATGACCGTCCCGTATTAAACAATGTGATAATTA 130
Db 19364 TGTCTCTGTCGGGTATTAAATGCAATGACCGTCCCGTATTAAACAATGTGATAATTA 19305
QY 131 CTCGGTTACCGGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGATCCCGCTTATCT 190
Db 19304 CTCGGTTACCGGAAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGATCCCGCTTATCT 19245
QY 191 GTTGCAATTTCCCTTACGCTGACTAGCCAGAGACACAATGATCTGTGCCCTTCTGTTAAT 250
Db 19244 GTTGCAATTTCCCTTACGCTGACTAGCCAGAGACACAATGATCTGTGCCCTTCTGTTAAT 19185
QY 251 ATCAAAACCGGTACTCAATATCTTCTGTGGCGCTGGCTGCCATCATCCGGAGCGTTCCGG 310
Db 19184 ATCAAAACCGGTACTCAATATCTTCTGTGGCGCTGGCTGCCATCATCCGGAGCGTTCCGG 19125
QY 311 TCGGGATAAAAAATTCGCGAGTGCAGCGGTGCCATGCAGACACATCCCCACGGGTAAACAG 370
Db 19124 TCGGGATAAAAAATTCGCGAGTGCAGCGGTGCCATGCAGACACATCCCCACGGGTAAACAG 19065
QY 371 CGTCCCTGTGCACATCTTCTGAAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 430
Db 19064 CGTCCCTGTGCACATCTTCTGAAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 19005
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTTCGGAAGG 490
Db 19004 GGGCAGCCGGAGACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTTCGGAAGG 18945
QY 491 CATGACACACCACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 550
Db 18944 CATGACACACCACTGGCAATCGGTAGACCGTAAACGCTGGGAAAGGACCTGCGCAT 18885
QY 551 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTTCGTAATTC 610
Db 18884 TAACACATCTCCGCTCATTTCCCAAGGTGTTCTGCTGACGAGACGCTTCGTAATTC 18825
QY 611 TTACAGCCGGGCCCCACACGAGCGAGCAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 18824 TTACAGCCGGGCCCCACACGAGCGAGCAATGATTTCCCTTCCATCTTCAGCTGATA 18765
QY 671 CAATACACGAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
Db 18764 CAATACACGAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 18705
QY 731 AAGCGGAACATCTGCTGTAGTACGAGCCAGCGGTGGATATGAGGGGTAAACGGTGCAT 790
Db 18704 AAGCGGAACATCTGCTGTAGTACGAGCCAGCGGTGGATATGAGGGGTAAACGGTGCAT 18645
QY 791 GGCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCTCT 850
Db 18644 GGCTTCATATGCAATCGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGCTCT 18585
QY 851 GACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGATAGGACCAACATGAATACACA 910
Db 18584 GACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGTAGGACCAACATGAATACACA 18525
QY 911 CAGATAATTCAGGAAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 970
Db 18524 CAGATAATTCAGGAAACGTTCTGCTTACGGGTGATGATAGGTTTTTGTCTGCAAT 18465
QY 971 AGTGAAGCGGTGACAGCATATCAGCGCTGCTGCTGCTATATTTACTGTCATGGCCACT 1030
Db 18464 AGTGAAGCGGTGACAGCATATCAGCGCTGCTGCTGCTATATTTACTGTCATGGCCACT 18405
```

Db 18464 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCGTATATATTAATCTGATGCGCACT 18405
 Qy 1031 ATGCAGATGACCATGACGATGAGTTAAATCCCGATAAATCCGTCGAAGTCTGAGGATGGA 1090
 Db 18404 ATGCAGATGACCATGACGATGAGTTAAATCCCGATAAATCCGTCGAAGTCTGAGGATGGA 18345
 Qy 1091 AGGAAGGTGAGGCTGTCTCTGAAAGGAATAAAGTGACATCATGCCCTCTTTTCTGCGC 1150
 Db 18344 AGGAAGGTGAGGCTGTCTCTGAAAGGAATAAAGTGACATCATGCCCTCTTTTCTGCGC 18285
 Qy 1151 TTCGGAGCAATTTTACTTTTCTCTGCGAG 1181
 Db 18284 TTCGGAGCAATTTTACTTTTCTCTGCGAG 18254

RESULT 6

AB011549/c 92721 bp DNA circular BCT 27-APR-1999
 DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.

AB011549
 VERSION ToxR-regulated lipoprotein; tagA.
 KEYWORDS Escherichia coli (strain:O157:H7, sub_strain:RIMD 0509952)
 SOURCE plasmid:pO157 DNA.

ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (sites)
 AUTHORS Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
 Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
 Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
 and Shinagawa,H.

TITLE Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
 enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
 outbreak

JOURNAL DNA Res. 5 (1), 1-9 (1998)

MEDLINE 98290340

REFERENCE 2 (bases 1 to 92721)

AUTHORS Makino,K.

TITLE Direct Submission

JOURNAL Submitted (24-FEB-1998) Kozo Makino, Research Institute for
 Microbial Diseases, Osaka University, Molecular Microbiology;
 Yamadaoka, 3-1, Suita, Osaka 562, Japan
 (E-mail:makino@bks01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
 Fax:81-6-879-8320)

COMMENT On Apr 20, 1999 this sequence version replaced gi:3336997.

FEATURES Location/Qualifiers

1..92721

/organism="Escherichia coli"

/strain="O157:H7"

/sub_strain="RIMD 0509952"

/db_xref="taxon:562"

/plasmid="pO157"

/note="RIMD 0509952 is a strain of enterohemorrhagic E.

coli, EHEC O157:H7"

join(92527..92721,1..2502)

/gene="tagA"

join(92527..92721,1..2502)

/gene="tagA"

/codon_start=1

/transl_table=11

/product="ToxR-regulated lipoprotein"

/protein_id="BAA31757.3"

/db_xref="GI:4666293"

/translation="MTKMERWTPMKLYLSTLAPLAGVSTADNNSALVF

NTSQPINDAGSLAAEVKFAQSQILPAHPKQSGSQHLTSRLSLILVRPKADKTP

VQVARDNNKILGTLTYLPSSLPDTIYHLDGPEGGIDFTPHNGTKKLINIAEVN

KLSDAGSSSHLTHNALVEIHTANGWRDYILPQGPDLGKMYRFVSSAGYSTV

FYGRKVTLSVNTLLFYVNGQWFRSGELENRTIYAQHSAPLPAHWVPLGLNV

IKQNLGRNLIDIKIGAPGELLHTIDIGMLTTPDRDFDFAKDKAHHRYFOFIPYSR

MTVNYAPLHLKVMYLPTELLTMDPGNGGWSGTMQRIGKELYSHGIDNANYGLN

STAGLGENSPHYVAQLAAHNSRGNYANGLOVHGSGGGGIIVLDSLGNFEFHEVGH

NYGLIGHYVDGFKGVSHNSAENNNSTWGDGDKKRFIPNFYPSQTNKSLNNOCQPEF

DKHKGFAMAGGSPFSANRFTWTYTPNSSALIORFFENKAVFDSRSTGFSKKNADT
 QMEPEYHTIDRAEQITASVNELSESKMAELMAEYAVYVKVHMWNGNWRNRYIPTASA
 DNRSGLITINHEAGYSYLVINGEKYSQKFSVSDGQKWEREDYVTRREARKPE
 QFQVPTLVLYGVDPEGLSSYIYPMATGYFTYSDSDONLSDNCOLOVLYREKQGL
 REFLANHRANTVMNKHFNPTESQPTQATLVCNNKILDTKSLTPAPEGLTYTVNGO
 ALPAKENGCCIVSNKSKRYCLPVGQSRGYSILPDWIVGQEVYVDSGAKAKVLLSDWN
 LSYNRIGSFVGNVNPNADMKVKAMNGQYLDPSKPSMRKRVYK"

gene

CDS

/gene="etpc"

/gene="etpc"

/codon_start=1

/transl_table=11

/product="Type II secretion pathway related protein"

/protein_id="BAA31758.1"

/db_xref="GI:3336999"

/translation="MLFSLSPGRDGLFKIDIVLKMLTPNRLCLVILLIAGYQLVSVI
 HFHWTQAASVPLSRVAPETATGDTQTEERFVTLFGRASPLSSEGRAGQETMPSLS
 DLLSGEDLDVRLGILYSVAEHSVAIFAHNRRQFSLSGEKVPSVDATISAIFSDHIV
 INYQKTVSLPLRYDNTTEKNAYDNNNLTVGDVITQDNFRVESVFDIMSFSAVTYNT
 LSGYRLIPGKHSSLYNAGLHNDLAVSVNGSELDRQAOQIMKOLPELKEIKTIVE
 RDGLYDAFIAVGEN"

gene

CDS

/gene="etpd"

/gene="etpd"

/codon_start=1

/transl_table=11

/product="Type II secretion pathway related protein"

/protein_id="BAA31759.1"

/db_xref="GI:3337000"

/translation="MLNEEQYYQFFLSVLDVGVFAVDMHNGILKVVRSKDAKTSAPV
 AASVSPGTGDEVTVRVVSVNAARDLAPLLRLQNDNAGAGSVHYEYPSNVLLMTGR
 VADMKRLMEIVRDKVGNRSVATVPLTASATDVARLVLTETKTDTATLPAWMTAK
 LYADERTNSLVSGEPISQORIISIIKOLDROEDVQGNTKVLYLAKAKOLDELVLGT
 ISSIENDSKSPSTEALRKGYTIKSHQTNALLITGAPDVIRDLNENVSQIDIRRPQ
 VLVEALIAIQDADGNIQGNVNRHAGVAQTGLPTTMTVQTRQNEILDSQDQNA
 LSMFNGIAGFYQGNWMLLTALSTSSKNLILATPSIVTLDMNEATFVNGEVPVLSG
 SQTSGDNIFNTVERKTVGILKRVKQINEGDVSLEIEQEVSGVADTAVATTDLGA
 TENVTRINAMLVNGETVVGGLDKSIRGSESKVPLIGDIPVGLHLPFRKSEOTAK'
 RNLMFIPTIIRERDGFHSAEKYQSEFNQVQSRGKETTALTTLNEQLRLSPDQD
 DPAFRKVAALAAFYAQEM"

gene

CDS

/gene="etpe"

/gene="etpe"

/codon_start=1

/transl_table=11

/product="Type II secretion pathway related protein"

/protein_id="BAA31760.1"

/db_xref="GI:3337001"

/translation="MSRVQNVSESRPLLPFSFRTQRIILLREQGNRVFCMEDTPA
 SALLFVRVAEGPLNVTTVSAEAFKQLVSSYORDSDEAROMMAEIGNEMDPTVAGE
 LPDREDLLDNDADAPILRLINAMLTAEKAKADJHETYERHLQVRFIDGVLREIL
 RLHRLNLSLIRIKVMARIDIAERVPQDGRVLRIGRAVDVVRVSTPSNNGRIV
 LRLLDKNSVSLDIALGALQQOQHIDALIRPHGIIIVTGTSGKSTLYALSL
 NPDNRNIMTVEDEVYELDGIISQTVNPKVDMTFAFRAILRQDPVLYVGEIRDE
 TQAIQVQASLTGHLVLTSTHNSAAGSLRQDMGIPPELSTSLAVLAQRLVLTLC
 PRCPQCVSTELAMDMDIPPTTIWOPAGCHCSTFTGVHGTGTGHEILLIDDRITA
 IYQGEGLGITRLACGRYTLRLGDRGRQVLAGETSWEEVVRVVTESRLQEE"

gene

CDS

/gene="etpf"

/gene="etpf"

/codon_start=1

/transl_table=11

/product="Type II secretion pathway related protein"

/protein_id="BAA31761.1"

/db_xref="GI:3337002"

/translation="NALFYQASDIHGRKRSGLLEADSARHQLRLREALIPVRLDE
 KOVHHKRSRLRFRPGSSAEALTLQATLVAASLPLEEALDALLROSEKPRQ
 RLIIAARVTRVLEHSLAAMGMFPGTPELVCAMVAGETSGRLDVLRLADYEQ
 RQIMRNLRLQALLPCVLTAVGVITAILLTAVTAVKVEQFIHMKQTLPLSLVLMGA

```

AVSQTWGPWLLLAALGGTAGRMILHQPSORLAFPHILLRLPVGGRISRGLNTARYA
RTLSTLNSAVPLLQAMHI SGLVLSNDWAHQDLAAELVREGVLSHQALEQTSLEPLF
MWRHMISGENSGELDSKMLERAADNQDRSFSTQMLAIGLFEPLLVGMAGVVLVFLIF
AIALPPLQLNNMNN"
gene
CDS
8193..8627
/gene="etpG"
8193..8627
/gene="etpG"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337003"
/translation="MRQHQRGFTLLIMVIVILGVLSLVVPLNLMKNKDKADROKV
MSDLVLESTLDMTRFLDNRRYPTTEQGRALVSKPTVQPEPRMTRODGIIRLPQDFW
GGDYOLLNPGQYSDIDIFSPGPDGVPNTEDDIGNWTLGNAQP"
8624..9178
/gene="etpH"
8624..9178
/gene="etpH"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337004"
/translation="MSORGFTFLBMLILLMGITAGLVLSMSPDSQAQNHLOQORREL
QAQILDYDRSQDDLLMGIOVREDGKFKVLQRGTAESPPTLAEGDIIWQGYVWQTW
OPRRAMGGKLPDDVLELQYLKGLQWMSHDDGAEPDILLPGGEVTFRLLRFRQVG
EAVVGVGGDNLMTLFEVEVSL"
9193..9540
/gene="etpI"
9193..9540
/gene="etpI"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31765.1"
/db_xref="GI:3337005"
/translation="MTLIEVVALVWFALAGMALMQASTQQAAGIGRMEEKVLAGWLA
DNOMVQLOLEKTPWENGKGTISFAGTEWTLRWGPDSDVPQPSRSLVEDTSAAME
SALVLSRSSVRE"
9537..10136
/gene="etpJ"
9537..10136
/gene="etpJ"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31766.1"
/db_xref="GI:3337006"
/translation="MSQORVKGFTLLEMLLALAVFAALSISAFVQLGGIIRAHELSDR
KVRLAEIQGISOMERDLOMTQPRHSGNEGILLAAAPHLKSDMGCISFTRSNLNP
AGMLPRPELQWVGYRLROOKLERLSYFHDHPGCVSPDVRMLDGVHAFRLRFVFGD
WQAKWDSGILPQAVEVTLVMDDFAEPLRLFLYSKETAE"
10133..11110
/gene="etpK"
10133..11110
/gene="etpK"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31767.1"
/db_xref="GI:3337007"
/translation="MKLREQGVALLVLLILSLMVTIAVIAERNGRFTLRTVAQLDQ
LQAKNGYTAETIAKQILORSRQSPKTHLAONWASOQEFETRGDVRGQIVDAQA
CFNLNAINGVVDLTSIYPYARISFQOILLINVELLQARQVTAALRDWIDRDKPVRG
GAEDVYMGPEPPIIPAAQPMQDYSERLNLINGIDARLYRKLLPTVCVLPSTLDSLVNN
TLLDSQAPLLALHFLTPKDSLPFLVTLQRRPRTGWESVAAPFLDAPPLKDDTSAAMPY
LAVSSNYFLVRLHRSSEHLFSQOITLQWREERFIQRQYGLTMRVP"
11473..12321
/gene="etpL"

```

Query Match

98.7%; Score 1166.2; DB 1: Length 92721;

Db 87919 ATGCAGATGACCATGAGTTCATTAATCCCGATAATCCGTCGAAGTCTGAGGATGGA 87860
QY 1091 AGAAGGTGAAGCTGTTCTCTGAAGCAATAAAGTGACATCATGCCCTTTTCTGGC 1150
Db 87859 AGAAGGTGAAGCTGTTCTCTGAAGCAATAAAGTGACATCATGCCCTTTTCTGGC 87800
QY 1151 TTCGGGAGCAATTTTACTTTTCTCTGCGAG 1181
Db 87799 TTCGGGAGCAATTTTACTTTTCTCTGCGAG 87769
RESULT 7
AX191725/c 92721 bp DNA linear PAT 15-AUG-2001
LOCUS AX191725 Sequence 7 from Patent WO0149775.
DEFINITION AX191725
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 92721)
AUTHORS Iversen, P. L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source Location/Qualifiers
1..92721
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 24870 a 20267 c 23868 g 23716 t
ORIGIN
Query Match 98.7%; Score 1166.2; DB 6; Length 92721;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAATATAAAAAATGCCCCATCCCGCGCTCCAGCTGAAAGTAGGCC 70
Db 88939 TGGAAAAAGCCAAATATAAAAAATGCCCCATCCCGCGCTCCAGCTGAAAGTAGGCC 88880
QY 71 TGTTCCTGCGGTATTAATGCAATGACCGTCCCGGTATTAACCAATGTGATAATTA 130
Db 88879 TGTTCCTGCGGTATTAATGCAATGACCGTCCCGGTATTAACCAATGTGATAATTA 88820
QY 131 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCCGCTTATCT 190
Db 88819 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCCGCTTATCT 88760
QY 191 GTTGCAATTCGCCCTTAGCCTGACTAGCCAGACACAAATGATGTCGCGCTTCTGTTAAT 250
Db 88759 GTTGCAATTCGCCCTTAGCCTGACTAGCCAGACACAAATGATGTCGCGCTTCTGTTAAT 88700
QY 251 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCTTCGG 310
Db 88699 ATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCTTCGG 88640
QY 311 TCGGATAAAAATCGCGAGTGCCTGATCGACACACATCCCGGAGGATGTAACAG 370
Db 88639 TCGGATAAAAATCGCGAGTGCCTGATCGACACACATCCCGGAGGATGTAACAG 88580
QY 371 CTTCCCTGTCAATTTCTTGAATGACATGAGGATCCCGGCTCTCAGTGGGATTAAC 430
Db 88579 CTTCCCTGTCAATTTCTTGAATGACATGAGGATCCCGGCTCTCAGTGGGATTAAC 88520
QY 431 GGGACCGCGGAGTGCCTGATCGACACACATCCCGGAGGATGTAACAG 490
Db 88519 GGGACCGCGGAGTGCCTGATCGACACACATCCCGGAGGATGTAACAG 88460
QY 491 CATGACACACATGGCAATCCGGTAGCCGGTAACTGGGAAAGGACCTGCCAT 550
Db 88459 CATGACACACATGGCAATCCGGTAGCCGGTAACTGGGAAAGGACCTGCCAT 88400

QY 551 TAACATATCCGCTCATTCCTCAGGTGTTCTCTGCTGACGACGACGTGCTTCGTATTC 610
Db 88399 TAACATATCCGCTCATTCCTCAGGTGTTCTCTGCTGACGACGACGTGCTTCGTATTC 88340
QY 611 TTCAGCCCGCGGCGCCACACAGGACGAGCAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 88339 TTCAGCCCGCGGCGCCACACAGGACGAGCAATGATTTCCCTTCCATCTTCAGCTGATA 88280
QY 671 CAATACACGACGACATAAATTCATGCTTTTTCGGGAGCTAGCATCCCACTGAAACGAT 730
Db 88279 CAATACACGACGACATAAATTCATGCTTTTTCGGGAGCTAGCATCCCACTGAAACGAT 88220
QY 731 AAGCGGAACATTTCTGCTGATGTCAGCCCGGCGGTGGATATGACGGGTAAACGCTGCGAT 790
Db 88219 AAGCGGAACATTTCTGCTGATGTCAGCCCGGCGGTGGATATGACGGGTAAACGCTGCGAT 88160
QY 791 GCGTTCAATTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850
Db 88159 GCGTTCAATTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88100
QY 851 GACACCTTCGCGCATCAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 910
Db 88099 GACACCTTCGCGCATCAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88040
QY 911 CAGATAAATTCAGGAAACGTTCTGCTGCTTACGGGTGATGAGGTTTTTTTGTCTGCAAT 970
Db 88039 CAGATAAATTCAGGAAACGTTCTGCTGCTTACGGGTGATGAGGTTTTTTTGTCTGCAAT 87980
QY 971 AGTGAAGCGGTGACACATATCAGCGCTCAGTCTGCTGCTATATCTGCTCATGGCCACT 1030
Db 87979 ACTGAAGCGGTGACACATATCAGCGCTCAGTCTGCTGCTATATCTGCTCATGGCCACT 87920
QY 1031 ATGCGAGATGACGACATCAGGTTTAAATCCCGGATAATCCGTCGAAGTCTGAGGATGA 1090
Db 87919 ATGCGAGATGACGACATCAGGTTTAAATCCCGGATAATCCGTCGAAGTCTGAGGATGA 87860
QY 1091 AGAAGGTGAAGCTGTTCTGAAAGCAATAAAGTGACATCATGCCCTTTTCTGGC 1150
Db 87859 AGAAGGTGAAGCTGTTCTGAAAGCAATAAAGTGACATCATGCCCTTTTCTGGC 87800
QY 1151 TTCGGGAGCAATTTTACTTTTCTCTGCGAG 1181
Db 87799 TTCGGGAGCAATTTTACTTTTCTCTGCGAG 87769

RESULT 8
AF043470 5612 bp DNA linear BCT 25-JUN-1998
LOCUS AF043470
DEFINITION Escherichia coli plasmid pO157 ecf4 gene, partial cds; and ecf3, ecf2, and ecf1 genes, complete cds.
ACCESSION AF043470
VERSION AF043470.1 GI:3253288
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 5612)
AUTHORS Boerlin, P., Chen, S., Collbourne, J.K., Johnson, R., De Grandis, S. and Gyles, C.
TITLE Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids and the locus for enterocyte effacement in shiga toxin-producing E. coli
JOURNAL Infect. Immun. 66 (6), 2553-2561 (1998)
MEDLINE 98261495
PUBMED 9596716
REFERENCE 2 (bases 1 to 5612)
AUTHORS Boerlin, P. and Gyles, C.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1998) Pathobiology, ON Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada
FEATURES Location/Qualifiers

source	1. .5612 /organism="Escherichia coli" /strain="4304-pr14" /serotype="O157:H7" /db_xref="taxon:562" /plasmid="p0157" /note="produces Shiga toxin" complement(<1. .705) /gene="ecf4" complement(<1. .705) /gene="ecf4" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24348.1" /db_xref="GI:3253289" /translation="MVPPSAVLCYHNEISROIQIPVNMKNIRTEFIPRNFNLTPFPYWM TWGIGICVFAMVPALRDLPLGLKMLVGRGKSARORALINLSLCPPEYSDRKE NIVDAMFATASMAVLMALSGDPKISHRIWNGLEIVKMAQNNKVIPLVPHAW GYDIPAMLAASGRKMAFMHQRNPVVDVYVNSVRRREGGKLHARDGIASFVRSR QGTYGILPDQDHGPEF" complement(751. .2499) /gene="ecf3" complement(751. .2499) /gene="ecf3" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24349.1" /db_xref="GI:3253290" /translation="MHLNTGQNPFTSWALGWAIFYFGFFSTLLQVILFSSGYSGTN GKDSLFSCMLDIPFLPDRIKIIAAVVGFILMGTSIAALCYIFLYGHEFSQVLF VMEFNARAGSFYSQFSLKLLILSVYAVSVFLWTLRPVYIPLWRRIVSFLL YALLHPVGLKSIROEPLNDLTKLASMEAPAPQWFSVSYQYHQOLNALTFLNE NSALPLGNLDESGRPRTLVLIGESTORMSLYGLRETTPELDALRKTDPGLT VFNVYASPYTEALQOALTANENKPDLYLTOPSLMMKQAGYKTFWITNQOIT ARTMLTVSRQTDROYVMNQRTQSAREDYNVLPFREVLNDPAKLLIIVHLLGT HIKYRYPGGGRFGTIGTIPGUNAKELEYINDYDANLFDHNVASLIDKDFRAT APDGLLYPDSGHEEYDTPPYKTQGRNEDNTPRMYTPVFLMTSEKWHAAHPRDFS QYDVRYSIAELIHTWMSLAGLYDGYDTPTRSLVNPQFRETTRWIGNPKYKKNGLTDFD LFRWVSIFLSSVPP" complement(2589. .3695) /gene="ecf2" complement(2589. .3695) /gene="ecf2" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24350.1" /db_xref="GI:3253291" /translation="MMKILFTSSSDIGGOELQALQMTALQKQHSVLACREKSKI APARRGHDTVPFPPRNSLHLPISILFRRIIGEFKPDVIVICHSGHDSNIAGLSLRIC CHRSIVROKTYITRKYTFSLNYLQDFIVPSSAMHLMABGVRTPTVPIPPGPDW PALHNGAMPLPHIHAWAASADNVLIVQVGMRLPEKGEHEMLRYLYOLKMEGKFR WLVVAGREYFASLQQOETHLGMGSDVLMAGALFPALPVRIASVVVWPSNEAFGM VLAASVSGVPVIASTGGIPDVQKNVTGTLPLPGVDVSAWNTGALRDLFSRPERFRM AASAREDIYRFINDINTAQIIVLSAQAQKCNR" complement(3695. .4516) /gene="ecf1" complement(3695. .4516) /gene="ecf1" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24351.1" /db_xref="GI:3253292" /translation="MLNARHLPVLMTHVHVSRCPLVTLSPETCEONKMLADNNRTV TSAELFYFGGFLPRKSVMLTFDDGYLDNWLRAWPLVKEYNLHAIHIFLTGLIGKN VRSRQEQYSHRCEYELIADNRSDEYVLRWSEYREMRDGLVEFHVHTSHKRWDRLS VSRACQRLMKEDILVQKCLTEKLGFCSSHLCPWEPGYNRYDINLAGKLGFSYLYTT ERMNCPENGLRIGRISTKERESGWLKRLFYITPLFSSVLALHKGPRLPDN" 1574 a 1399 c 1207 g 1432 t	Query Match 98.5%; Score 1163; DB 1; Length 5612; Best Local Similarity 99.6%; Pred. No. 0; Matches 1166; Conservative 0; Mismatches 5; Indels 0; Gaps 0; QY 11 TGGAAAAAGCCAAAAATAAAATGCCCATCCAGCGCGTCCAGCTGAAAGTAGGCC 70 Db 2415 TGGAAAAAGCCAAAAATAAAATGCCCATCCAGCGCGTCCAGCTGAAAGTAGGCC 2474 QY 71 TGTTCGTCCGGTATTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAAATTA 130 Db 2475 TGTTCGTCCGGTATTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAAATTA 2534 QY 131 CTCCTGTACCGGAAACCGCTGAACAAAATTGCGGTGAAAGAGAGATCGCGGTATCT 190 Db 2535 CTCCTGTACCGGAAACCGCTGAACAAAATTGCGGTGAAAGAGAGATCGCGGTATCT 2594 QY 191 GTTGCAATTTCCCTTACCTGACTAGCCAGAGACACAATGATCTGTGCGTCTCTGTTAAT 250 Db 2595 GTTGCAATTTCCCTTACCTGACTAGCCAGAGACACAATGATCTGTGCGTCTCTGTTAAT 2654 QY 251 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310 Db 2655 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 2714 QY 311 TCGGGATAAAAAATCGCGAGTGGCGGTCCATGCAGACACATCCCCACAGGGGTACAG 370 Db 2715 TCGGGATAAAAAATCGCGAGTGGCGGTCCATGCAGACACATCCCCACAGGGGTACAG 2774 QY 371 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430 Db 2775 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 2834 QY 431 GGCACCGCGGAGACTGAGCGTTTCCAGCCAGTACCATACCAACAGCTTCATTTCCGAAGG 490 Db 2835 GGCACCGCGGAGACTGAGCGTTTCCAGCCAGTACCATACCAACAGCTTCATTTCCGAAGG 2894 QY 491 CATGACACACACTGGCAATCCGGTAGACCGGTAAACGCTGGGAAAGGACCGTCCCAT 550 Db 2895 CATGACACACACTGGCAATCCGGTAGACCGGTAAACGCTGGGAAAGGACCGTCCCAT 2954 QY 551 TAACATCTCCGCTCATTTCCAGGTGTCTGCTGACGAGAGTCTCTCGTATTC 610 Db 2955 TAACATCTCCGCTCATTTCCAGGTGTCTGCTGACGAGAGTCTCTCGTATTC 3014 QY 611 TTCACGCGCGCGCCACCACGAGCCAGGAAATGATTTCCCTTCCATCTTCAGCTGATA 670 Db 3015 TTCACGCGCGCGCCACCACGAGCCAGGAAATGATTTCCCTTCCATCTTCAGCTGATA 3074 QY 671 CAATACACGAGCATAAATTCATGTCTTTTCGGGACGTAGCATCCCCACCTGAAGCAT 730 Db 3075 CAATACACGAGCATAAATTCATGTCTTTTCGGGACGTAGCATCCCCACCTGAAGCAT 3134 QY 731 AAGCGAAACATTTCTGCTGATGCAGCCAGCGTGTGATATGAGGGGTAAACGGTCCCAT 790 Db 3135 AAGCGAAACATTTCTGCTGATGCAGCCAGCGTGTGATATGAGGGGTAAACGGTCCCAT 3194 QY 791 GGCTTCATTTATGCAATCGCGCGCGAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGCTCT 850 Db 3195 GGCTTCATTTATGCAATCGCGCGCGAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGCTCT 3254 QY 851 GACACCTTCGGCATCAGATGCGCCATCATGGGTGAGATAGGCACAAACAATGAATCACA 910 Db 3255 GACACCTTCGGCATCAGATGCGCCATCATGGGTGAGATAGGCACAAACAATGAATCACA 3314 QY 911 CAGATATTCAGGAAACCGTTCTGCTTACGGGTGATGTAGGTTTTTTTGTCTGCAAT 970 Db 3315 CAGATATTCAGGAAACCGTTCTGCTTACGGGTGATGTAGGTTTTTTTGTCTGCAAT 3374 QY 971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCCTGCTATATTACTGTGATGGCACT 1030 Db 3375 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCCTGCTATATTACTGTGATGGCACT 3434
--------	---	--


```

/codon_start=1
/transl_table=11
/product="ospf, secreted by the Mxi-Spa secretion
machinery, function unknown"
/protein_id="CAC05773.1"
/db_xref="GI:12329042"
/translation="MPIKPKLNLDSLNVKSETPOMLSANERLKNFNILYNOIR
QPAYPKVASNPVTSDICQPFVVMQGFQIVNHSQDFVHACRENPSKQDFVQDK
FHSIAREQVPLAFQILSGLLFSESDPIDKWKITDMNRVSOQSRVIGIAQFTLYVXSD
QECQSALLHRIQFIMCLNESNLLRSKAPGEYPASDVRPDMKIVSYRNEIHDR
DSEROEQLREPFYRLMIE"
/complement(13393..13971)
/notes="orf13, length= 192 aa, unknown"
/codon_start=1
/transl_table=11
/protein_id="CAC05774.1"
/db_xref="GI:12329043"
/translation="MKVSPKSLGYIFHDYINKKHTIDFNDVVRKAVLSKINELNAC
HKVAFLEAKDNHETTKDKAKIIDTLTENYSIEFQOLMNISETLNSLYITPESGF
VSFVNEGKICHTAYVKSSDNMTYHANGSSIDKXITDMCGLICMRHIESIGLIIFM
LDRKVLASAEFENRNGWRRAFCANLYKCV"
14014..14241
/notes="IS2.02, 90% id over 228 nt with IS2, from 4 to 230"
/complement(14367..14629)
/notes="IS21.01, 97% id over 263 nt with IS21, from 1825 to
2098"
/misc_feature
complement(14630..15558)
/notes="IS91.01, 97% id over 929 nt with IS91, from 191 to
0"
15559..15924
/notes="ISshf5.02a, id to ISshf5 from 1 to 366"
/complement(15925..17350)
/notes="IS4.01, 99% id over 1426 nt with IS4, from 1 to 0"
17351..19039
/notes="ISshf5.02b, id over 1689 nt with ISshf5, from 335
to 0"
19326..19523
/notes="IS3.01, 88% id over 198 nt with IS3, from 1062 to
0"
20459..20601
/notes="IS630.02, 84% id over 143 nt with IS630, from 999
to 1141"
20964..21641
/gene="ospd1"
CDS
20964..21641
/gene="ospd1"
/notes="ORF21, length = 225 aa, 34.3% id in 140 aa overlap
with FCSHET2B.1 (S.flexneri SenA - 565 aa)"
/codon_start=1
/transl_table=11
/product="ospd1, secreted by the Mxi-Spa secretion
machinery, function unknown"
/protein_id="CAC05775.1"
/db_xref="GI:12329044"
/translation="MSINNVGLHPANNKWHLLIGSNTANENKMKNNIINVTALIS
HAINRKSQGVSGVFRKLAKTIONISIPKNNKYNRNLFLSLIWHGNADARKYSE
SLUAEIPEERLEVLAAARNAGESALFIALQEGHSAAYQAGDFIKTFLDLSKETIK
LLDVRNDEGLPLFLAAGRGNIEMMAYINICHHSIGIKLITEIADRLNNNEQDMFNIS
DKIQELF"
21977..22105
/notes="IS3.02, 89% id over 129 nt with IS3, from 1130 to
0"
22153..22467
/notes="ORF22, length = 104 aa, unknown"
/codon_start=1
/transl_table=11
/protein_id="CAC05776.1"
/db_xref="GI:12329045"
/translation="MLQRQKRGVFAQLPVDFVAIEPDSVQGVKGRANLTNRCFIIRI
27.8%; Score 328.6; DB 1; Length 213494;
Query Match
Best Local Similarity 59.3%; Pred. No. 1.4e-94;
Matches 578; Conservative 0; Mismatches 394; Indels 3; Gaps 1;

```

RESULT 10

AF386526/c

LOCUS

AF386526

221618 bp

DNA

circular

BCT 01-FEB-2002

```

DEFINITION Shigella flexneri 2a strain 301 virulence plasmid pCP301, complete
sequence.
ACCESSION AF386526
VERSION AF386526.1 GI:18462515
KEYWORDS Shigella flexneri 2a.
SOURCE Shigella flexneri 2a
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Shigella.
REFERENCE 1 (bases 1 to 221618)
AUTHORS Jin,Q., Zhang,J.Y., Liu,H., Yang,J., Yang,F., Zhang,X.B.,
Wang,J.H., Yang,G.W., Wu,H.T., Dong,J., Sun,L.L., Xue,Y.,
Zhao,A.L., Gao,F.S., Zhu,J.P., Kan,B., Chen,S.X., Yao,Z.J.,
He,B.K., Chen,R.S., Ma,D.L., Yuan,Z.H., Xu,J.G., Wang,Y., Shen,Y.,
Lu,W.C., Qiang,B.Q., Wen,Y.M. and Hou,Y.D.
TITLE Complete DNA sequence and analysis of the large virulence plasmid
pCP301 of Shigella flexneri
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 221618)
AUTHORS Jin,Q., Zhang,J.Y., Liu,H., Yang,J., Yang,F., Zhang,X.B.,
Wang,J.H., Yang,G.W., Wu,H.T., Dong,J., Sun,L.L., Xue,Y.,
Zhao,A.L., Gao,F.S., Zhu,J.P., Kan,B., Chen,S.X., Yao,Z.J.,
He,B.K., Chen,R.S., Ma,D.L., Yuan,Z.H., Xu,J.G., Wang,Y., Shen,Y.,
Lu,W.C., Qiang,B.Q., Wen,Y.M. and Hou,Y.D.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2001) Microbial Genome Center of Chinese Ministry
of Public Health, 100 YingXin Street, Xuanwu, Beijing 100052, P. R.
China
FEATURES
    source      Location/Qualifiers
                1. .221618
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /note="virulence plasmid"
    source      26244..26633
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="Inse"
                /note="virulence plasmid"
    source      complement(32084..32752)
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
    source      43481..44149
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
    source      50808..51476
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
    source      60987..61271
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
    source      60987..61271
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="ISO-IS1N"
                /note="virulence plasmid"
                /gene="CP0001"
                /complement(191..493)
                /gene="CP0001"
                /note="similar to a 301 aa protein from bacteriophage Sfx
                gb: AAD10297.1"

```

```

    source      /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
                70195..71403
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="tn10"
                /note="virulence plasmid"
                complement(72090..72881)
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
                complement(75050..76258)
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="tn10"
                /note="virulence plasmid"
                138916..139182
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
                143657..144064
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
                154128..154796
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
                167032..167685
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS1328"
                /note="virulence plasmid"
                complement(172967..173239)
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="ISO-IS1N"
                /note="virulence plasmid"
                186968..187636
                /organism="Shigella flexneri 2a"
                /strain="301"
                /db_xref="taxon:42897"
                /plasmid="pCP301"
                /insertion_seq="IS629"
                /note="virulence plasmid"
                complement(191..493)
                /gene="CP0001"
                /complement(191..493)
                /gene="CP0001"
                /note="similar to a 301 aa protein from bacteriophage Sfx
                gb: AAD10297.1"

```



```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL72473.1"
/db_xref="GI:18462701"
/translation="MVHATRLMKHASSPGCWDFVEPKNTAVRSPESNRIAKSPVKTIK
CDYISIMPRKPDGLTAANKLAEPHEYNHHPHSDYRSFVRLQRANDRCLEI"
485..949
/gene="CP00002"
/gene="CP00002"
485..949
/note="similar to a 333 aa protein from Sphingomonas sp.
gb: AAD45420.1"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL72474.1"
/db_xref="GI:18462702"
/translation="MNHPSVTSLETPEARCSGVPLPACRORESIYGLIELFIOIV
HRLSVRSERLVLTLADQVHGKTHALLPRIAEALNPNPDGLYKVVVHLHIVEPDR
SGRRSSYLQNLRDVSGARDVNRKTRLPQDSGLFALVSDPLPRMISTVL"
1176..2042
/gene="ospB"
/note="ospB"
1176..2042
/gene="ospB"
/note="similar to a 288 aa protein from Shigella flexneri
emb: CAC05770.1"
/codon_start=1
/transl_table=11
/product="ospB"
/protein_id="AAL72323.1"
/db_xref="GI:18462551"
/translation="MLDGVRYPCRIVKNKNEISIDIAFAHIIKRVKNSCTHPKAA
VFLGEFGDSDNVLISMGQIIPRVFNKMLDYVFNKESKNDFLKMAESWLPQSEP
IVNNDDILALNAEVSVKAKIKITVNDIDFKEYNKVYILGHGSPGSHOLGISELID
VORTIISRMKDCGILNKVDIREFTSCGSADKVPKFNENPAESIISLINSIPFKFES
LJROIKKHLENDLSLGLKISGHYGVHYGQELFPYSHYRSISIPADPEHVKRSS
OKRTFINNELD"
2371..3111
/gene="apy"
/note="CP00004"
2371..3111
/gene="apy"
/note="similar to a 246 aa protein from Shigella flexneri
gb: AAL21206.1"
/codon_start=1
/transl_table=11
/product="apyrase"
/protein_id="AAL72358.1"
/db_xref="GI:18462586"
/translation="MKTKNFLFLCIATNMIFIPSANALKAEGFLTOOTSPDSLILPP
PPAENSIVVFOADKAHYEFGRSRDLNRRVLASEDAYYENFGLAFSDAYGMDISRENTP
```

```

Query Match      27.88; Score 328.6; DB 1; Length 221618;
Best Local Similarity 59.38; Pred. No. 1.4e-94;
Matches 578; Conservative 0; Mismatches 394; Indels 3; Gaps 1;

QY 210 TGAATGACGACGAGACAAATGATCTGTGCCGCTTCTGTTAATATCAACACCGGTACTCAATA 269
DB 193635 TGCTTCGCTAATGTAGTATTTTAAACGCGAGTTTATTAATATCAACACCGCTCTCTATA 193576

QY 270 TCTTCTGCGGTGGTGGCCATCATCCGGAAGCGTTCGCGTGGGATATAAAATCGCGC 329
DB 193575 TCCTGTTTTCCTGGCGAGCATCTGATAAAACGCCGCGGTCAATTAATAAAATCATTC 193516

QY 330 AGTGGCGCGTCCATCGACACATCCCCACCGGTAAACGCGTCCCTGTCACTTCTTC 389
DB 193515 AGGCGGCATCATCCATCGCTTATTACCTGCTGGTAACAAATGCGCGGTCTGGTTATTC 193456

QY 390 TGAATGACATCAGGAGTCCCGCGCTCTCACTGCGGATACAGGCGCGGAGACTTGAC 449
DB 193455 TGAATACATCAGGAGTCCACCAATCTGACTGCCAGTACAGGCACAGAAATGCCGAT 193396
```

```

QY 450 GCTTACGCCAGTACCATACCAACAGCTTCAATTTTCGAAAGGCATGACACACCACTGGCA 509
DB 193395 GCTTTCGCCAGTACCATACCAACAGATTCGTTTCTGAAGCGAGAACCACCACTGGCA 193336

QY 510 ATCCGGTAGACCGGTACGCTGGGAAAGGCCACCTGCCATTAACACATCTCCGCTCATTT 569
DB 193335 ACCCGATATACGGGGCGGCGGAGAAACAAATGTGACGAATAAAACATCATCATGATG 193276

QY 570 CCCAGGTGTTCTGTCTGCTGACGACAGAGTCTGCTTCTGATTTCTCACCCCGGGGCCACCC 629
DB 193275 CCCATCTGTCAATCTGATACTGTAATGCTCCCGGAGTTCAGGCGAACCAGACCCACA 193216

QY 630 ACAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATACATACACGAGCAATAAAT 689
DB 193215 ATCAACACACAGAACTGTCTGCTCATTTTAAATGGAACAGTAATTAACATAAAT 193156

QY 690 TCATGTCTTTTTCGGGACGTAGCATCCCACTTGAACGATAGGCGGAACATTTCTGTGCT 749
DB 193155 TCGTGCCCTTTTTCGGGCGCAACATTCCTACCTGAGCAATAACGGGGGCATCCCTTCG 193096

QY 750 GATCGACGCCAGCGGTGGATATGACGGGTAAAGGTGCGCATGCTTCAATGCAATGCG 809
DB 193095 GACGCCAGCCCAAGAAAGAACATTTGGAGGCAAACTGTTTCGGGATCAACGTATATTC 193036

QY 810 GGCAGTGTGAACCCGGTGAATAACCGTTTACCGGTGTCCTGACACCTTCGCCCATCAGA 869
DB 193035 TGGAAATCAAGGCTGGAGGCACAACTAGTAAACCGGGTTCGACATCTCTCTGCTCCGA 192976

QY 870 TCGGCCATCATGGGTGAGATAGGCACAACTGAATCAACACAGATAATTCAGGGAAC 929
DB 192975 TGTGCTTCACTTGTTCGGGAACAACTACCTCATCGCAAAATGATTTATTTGAAAA 192916

QY 930 GTTCTGTCTTACGGGTGATGATAGTGTGTTTGTCTGACAATAGTGAAGCGGTG---ACAG 986
DB 192915 ACTTTGTTTTTTCGTGCAATATGCTTTTTCGCTGATAATCTGGAATGATGTTTCTTA 192856

QY 987 CATATCAGACGGTCTAGTCTCTGTATATATCTGATGCGCATATGCGAGATGACACAGA 1046
DB 192855 GTGAATAACCGGTACTTAACCAACAATATGCTATCATGCCCCACTGTGACAAAGCATTCGA 192796

QY 1047 TCAGGTTTAAATCCCGGTAATCCGTCGAAGTCTGAGGATGAAGGAAGTGAAGGCTG 1106
DB 192795 TTGGGCTGAACACCATGAACTATTCGAGTAATTCCTCATGAGTAGGAGTGTGAAGACTG 192736

QY 1107 TTCTTGAAGGAATAAAGTGACATCATGCGCTCTTTTCTGCTTCCGGAGCAATTTTA 1166
DB 192735 TTTTCGAAATAACCGGAATGTGATATCAATCCCAATTTACTGGCTTCAAAAGCAATTTG 192676

QY 1167 CTTTTTCTCTGCAG 1181
DB 192675 CTGTTTTCTCTGCAG 192661
```

```

RESULT 11
AF348706/c 221851 bp DNA circular BCT 24-APR-2001
LOCUS Shigella flexneri large virulence plasmid pWR501, complete plasmid
DEFINITION sequence.
ACCESSION AF348706
VERSION AF348706.1 GI:13310487
KEYWORDS Shigella flexneri.
SOURCE Shigella flexneri.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Shigella.
REFERENCE 1 (bases 1 to 221851)
AUTHORS Venkatesan,M., Goldberg,M.B., Rose,D.J., Grotbeck,E.J.,
Burland,V. and Blattner,F.R.
TITLE Complete DNA sequence and analysis of the large virulence plasmid
of Shigella flexneri
JOURNAL Infect. Immun. 69 (5), 3271-3285 (2001)
MEDLINE 21189246
```

PUBMED 11292750
REFERENCE 2 (bases 1 to 221851)
AUTHORS Venkatesan,M.M., Goldberg,M.B., Rose,D.J., Grotbeck,E.J.,
Burland,V. and Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2001) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. .221851
/organism="Shigella flexneri"
/db_xref="taxon:623"
/plasmid="virulence plasmid pWR501"
complement(191..493)
/gene="S0001"
complement(191..493)
/gene="S0001"
/note="residues 20 to 100 of 100 are 88.23 pct identical
to residues 217 to 301 of product encoded by
GenBank Accession Number AE000291 IS2 hypothetical protein
[Escherichia coli]"
/codon_start=1
/transl_table=11
/product="IS2 orf2, fragment"
/protein_id="AAK18312.1"
/db_xref="GI:13310488"
/translation="MVAHTELMKHAASPOCWDFVEPKNTAVRSPESNRIRAKSFVKTIK
CDYISIMPKDGLTAARKLAFAFHYNEWHPHSALDYRSPREYLQRANDNRCLT"
485..949
/gene="S0002"
485..949
/gene="S0002"
/note="residues 13 to 93 of 154 are 40.74 pct identical to
residues 5 to 85 of 333 of product encoded by GenBank
Accession Number U65001 putative resolvase [Sphingomonas
sp.]"
/codon_start=1
/transl_table=11
/product="putative resolvase, fragment"
/protein_id="AAK18313.1"
/db_xref="GI:13310489"
/translation="MNHPSVTSLETPEARCRSGVPLPACRORESIYGLIELFIQIV
HRLSVRSERLVLTADFORGHGKTALLPRIAFAALNPDGLVKEVYHLHIVEPDR
SGRSSYLAQLRDVSNRGDAVNGRTLPEQDSGLPALYSDPGLPRMISTVL"
1176..2042
/gene="S0003"
1176..2042
/gene="S0003"
/note="residues 130 to 273 of 288 are 28.57 pct identical
to residues 12 to 139 of 340 of product encoded by GenBank
Accession Number AL139075 ketoI-acid reductoisomerase
[Campylobacter jejuni]"
/codon_start=1
/transl_table=11
/product="orf, hypothetical"
/protein_id="AAK18314.1"
/db_xref="GI:13310490"
/translation="MNLGVRPYCRIVNKKNESISDIAFAHIIKRVKNSCTHPKRAAL
VFKEGFCDSNDVLSMGQOIPRFVKMLDYVFEKNEKSKNDFLKMAESWLPQSEP
IVINDDALNAAATFSVKAKIKTVDNDFEYKNVYIILGHSGPGSHOGLGSELID
VORIISPMKDCGILNKDIFRTSCGSADKVAPEKNNAESILSCILNISLPEPKES
LLSQIKKHLNDESLDGLKISGYHGYVHYGQELFPYSHYRSTSI PADPEHTVKKRS
QKKTFIINKELD"
2371..3111
/gene="S0004"
2371..3111
/gene="S0004"
/note="residues 1 to 246 of 246 are 99.18 pct identical to
residues 1 to 246 of product encoded by GenBank
Accession Number U04539 ATP diphosphohydrolase, apyrase
[Shigella flexneri]"
/codon_start=1
/transl_table=11
/product="apyrase"

gene
CDS
/protein_id="AAK18315.1"
/db_xref="GI:13310491"
/translation="MKTNLFILFIATNMTIPISANALKAEGLTQOTSPDLSILPP
PAAEDSVFLADKAHVEGRSLRDARVRLASEDAYENFGLASDAYGMDISRENTP
ILQLLTQVLDSDHYAYRNKAEYKRVRPVVIYKDATCTPDKDEKMAITGSIYSGHA
SFGWAVAILAEINQRAEILRGFERGESRVICGAHQWSDVEAGRLMGASVVAVLH
NTPFTKLSLEAKKEFEELNPTNELTP"
3272..3526
/gene="S0005"
3272..3526
/gene="S0005"
/note="residues 27 to 59 of 84 are 36.36 pct identical to
residues 427 to 459 of 729 of product encoded by GenBank
Accession Number D63999 hypothetical protein
[Synechocystis sp.]"
/codon_start=1
/transl_table=11
/product="orf, hypothetical"
/protein_id="AAK18316.1"
/db_xref="GI:13310492"
/translation="MLPSETMIWOPEFTDKTSRKLGAVPFTTCNVVLOGNGLPIPVV
DOYRNDNFRFRAQPKYILGHLSNRLPDTAFENKKNHF"
complement(3470..4441)
/gene="S0006"
complement(3470..4441)
/gene="S0006"
/note="residues 8 to 285 of 323 are 22.56 pct identical to
residues 56 to 301 of 612 of product encoded by GenBank
Accession Number AL008970], Hypothetical protein, len: 613
aa [Plasmodium falciparum]"
/codon_start=1
/transl_table=11
/product="orf, hypothetical"
/protein_id="AAK18317.1"
/db_xref="GI:13310493"
/translation="MIKSLIKYKHRIHTLILKLSLTTILFQGVESINHOEKLPLN
KTHTVYFGANAYIIDHSDPYGYMTLTFEHDNAIPPFYHEHQSFDFDNFKEVVDVDS
RYVHGQGTDPVIFNTDMRIGLIGLHIDFISKSDQGFREFCYNNKIDVSLDRII
NFVQLEYHIPMLSTDNFKIKLRDILEDAIKASNYEENNNKYTDKKAHQALAYS
LGDKADIALYLLSKNFETKQDAEMKMNINNYCNLYDVEYLLSKGANKVLEYFI
NGLYDVNKKFKOKANGDTMLDNAMSKDSKMDIFFIKKWSGIRQT1"
complement(4344..4742)
/gene="S0007"
complement(4344..4742)
/gene="S0007"
/note="residues 38 to 128 of 132 are 23.15 pct identical
to residues 41 to 123 of 839 of product encoded by GenBank
Accession Number 299115 yonO [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="orf, hypothetical"
/protein_id="AAK18318.1"
/db_xref="GI:13310494"
/translation="MKNFLRSIAQSYKMFSGTSPKSLNLSLEAPSGARSFRSL
EHLDRVSRHYISEIIQKVHPUSDRHLSIINFNFRHQSNSLNLSNINLKSFSD
KIQSNIOHKNYSEDIKEISNHDVFWG"
complement(5261..5812)
/gene="S0008"
complement(5261..5812)
/gene="S0008"
/note="residues 5 to 177 of 183 are 94.21 pct identical to
residues 42 to 214 of 523 of product encoded by GenBank
Accession Number AB024946 orf51 [Escherichia coli]"
/codon_start=1
/transl_table=11
/product="ISEC8 orf, fragment"
/protein_id="AAK18319.1"
/db_xref="GI:13310495"
/translation="MSRPEINRLKALVAKLQRMQFGKSSSEKRAKTERWIQEAQERI
SALOQEMAEITGEQDVPVLPSSLROSSARKPLPASLRPRVIRPEECCPACGELS
PLGCDVSEOLELISSAFVIEKORPKACRCDHIVQAPVPSKPIARSYAGAGLLAH
VTGKYADHILPIYRQSDLLFHTAI"
complement(5742..6164)

gene

[illegible]

Query Match	27.8%;	Score 328.6;	DB 1;	Length 221851;
Best Local Similarity	59.3%;	Pred. No. 1.4e-94;		
Matches 578;	Conservative 0;	Mismatches 394;	Indels 3;	Gaps 1;
QY	210	TGACTAGCCAGACAGACAATGATCTGTCGCCCTCTGTTAAATATCAAAACGGTACTCAATA	269	
DB	184983	TGCTTCGCCTAAATGTGAGTATTTTAAACGACGTTTATTAATATCAAAACCGCTCTCTATA	184924	
QY	270	TCTTCTCTCGCGCTGCGTCCCATCATCGGAAGCGTTCGCGTCGGGATAAAAAATCGCG	329	
DB	184923	TCTCTGTTTTGCGCTGGGAGCCATCTGATAAAACGCCAGGGTCATTAAAAAAATCATTC	184864	
QY	330	AGTGCSCGGTCCATCGACACACATCCCCCACGGGTAAACAGCGTCCCTGTGCATTTCTTC	389	
DB	184863	AGGGCGCACATCCATCGCGTCTATTACCTGCTGTTAAACAATGTCGCCGTTCTGTTTATTC	184804	
QY	390	TGAATCAGATCAGGGATCCCGCCCTCTCATCTGGGGATAACGGGCACCGCGGACATGAC	449	
DB	184803	TGAATAACATCAGGGATTCCACCAATCTGACTTGGCCAGTACAGGCACAGAAAAATGCCGAT	184744	
QY	450	GCPTTACGCCAGTACCATACCAAAACCTTTCATTTTCCGAAGGCATGACCACACACTCGCA	509	
DB	184743	GCPTTGCACAGTACCATACCAAAACGATTCGTTTTCTGAAGGCAGAACCCAGACTGGCA	184684	
QY	510	ATCCGCTAGACCGGTAAACGCTGGGAAAAAGGCGACTGCGCATTAACACATCTCCCGCTCATTT	569	
DB	184683	ACCCGATATACGGGGCGGCAGAAAAACATTGTCAAGCAATAAAACATCATCATGTCATG	184624	
QY	570	CCACAGTGTTCGTCTGCTGACGACAGCTGCTTGTTATTTCTTCAAGCCCGGCGCCACCC	629	
DB	184623	CCCATACTCTCAATCTGATFACTGTAAATGCTCCCGCAGTTTCAGGCGAACACAGACCCACAC	184566	
QY	630	ACGAGCCAGGAATGATTTCCCTTTCATCTTCAGCTGATACAATACACGAGCATTAAT	689	
DB	184563	ATCAACCCACAGACTGTGCTCCATCTCAATTTTAAATGGACACTGTAATTTCAACATAAAT	184504	

```

/gene="shf"
/notes="hypothetical; similar to Shigella flexneri Shf"
/codon_start=1
/transl_table=11
/product="Shf"
/protein_id="AAD34404.1"
/db_xref="GI:4959584"
/transl_table=11
/translation="MLNGGGILFKANHPVLMYHVSHPGLVLTSPETFRKQMKWLA
ENWKTLSSELEFFYRGKLPKSVMLPFDGVLNDWFOVPLNFEENLKAHVFLIT
SFINGVRHSPPKREYSHRDCHEQIATGNADNVMLESEVNEMLQSLVEFHVHYTH
TRWDKFTSRECKHLRDLGKREYLSKHLCKWEGYKNDYIQIAEELG
FHYLYTTERMNAPAKAARIGRISKERESCAWLKRRFLFYTPPFSSLLALYKGR
LPDD"
1403..2224
/gene="capu"
/notes="capu"
1403..2224
/gene="capu"
/notes="capu; hypothetical"
/codon_start=1
/transl_table=11
/product="hexosyltransferase homolog"
/protein_id="AAD34405.1"
/db_xref="GI:4959585"
/translation="MNILFTESSPNTGGQELQVAOMKALKMGHSHVLLVCRENSKIA
FEASKGIDITFALFRNSHIPVWRLGIVHSFQDAIVCHSGHDSNIVGLVRFETW
KHPRIIRKTYLTKTKFFSINHCFDEVIPGTSMTKHLDEQECRTVTVVPPGDF
QELVDRNSLPPSLVSWLASRRGCPVIAQVGMRLPEKGGDPLCWLTVFGFPELREH
LOYOISMGHDDVFIADNVLPPIYGLPVMWFCLQKTNLLVWCQKRRHFLCL"
2453..3403
/gene="VirK"
/notes="VirK"
2453..3403
/gene="VirK"
/notes="similar to Shigella flexneri VirK"
/codon_start=1
/transl_table=11
/product="VirK"
/protein_id="AAD34406.1"
/db_xref="GI:4959586"
/translation="MFSISNLSFGLKRIIVFSDSLPGKWEHRKFRFMYILRCSINP
VVSIRYVELRSLPCIEDILAIHPTLPARHPYLHKGRWTRGVILHYRFFVONL
PEKYSKFLPKQSVLQFQKNGENFDIOCSFGDFREGELMLSFYKNTVIRLARTY
SVILTONGHIAFGLGQAPKNTGPDVIRCATACVGLPKRIIFFAFCALMKACNVS
BCLAVSHSVFROKLTWYQKRTFVAVYSDFDVSVGKTCGDWYKLPQTVIRKPLTD
IPIKRSYRKQYALLDYIHTTIRSLIDAYPVHSHKQNSN"
BASE COUNT 993 a 668 c 800 g 1039 t
ORIGIN

Query Match
Best Local Similarity 20.9%; Score 246.6; DB 1; Length 3500;
Matches 550; Conservative 0; Mismatches 379; Indels 46; Gaps 3;

QY 210 TGACTAGCAGACACAAATGATCTGTCGGCTTCTGTTAATATCAAAACCGGTACTCAATA 269
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2441 TGCTTCGCTAATGTGAGTATTTTAAACACAGTTTATTAATATCAAAACCGCTCTCTATA 2382
QY 270 TCTTCTGCGCTGGCTGCCATCATCCGAGAGCTTCGCTCGGATATAAAATCCGGC 329
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2381 TCCTGTTTGGCAGCGGACCATCTGTATAAAACCGCCAGGCTCATTAATAAAATCATTC 2322
QY 330 ACTGCGCGGTCATATGACAGACATCCGCCAGGTAAACGCTCCCTGTGCACATTTCTC 389
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2321 AGGGCGCACATCCATCGCTGTTTATCTGCTGTACATATGTCGGCTGTGTTGTTTC 2262
QY 390 TGAATGACATCAGGATCCCGCGCTCTCATCTGGGATPAACGGCAGCCGCGGAGACTGAC 449
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2261 TGAATAAATCAGGATTTCCACCAATCTGAGTGGCCACTACAGGCAGACAGAAATCCGGAT 2202
QY 450 GCTTACGAGTACATACCAACCTTCATTTCCGAAGCATCACCACACACTGSCA 509
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2201 GCTTTCGCGACCACTACCAAGATTCGTTTTCTGAAGGCAGAACCCACGACTGCA 2142
QY 510 ATCCGGTAGACCGGTAAAGCTGGGAAAGGCGACCTGCCATTAACACATCTCCGCTCAT 569
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

2141 ACCGATATA -CGGGGCGGCAGGAAAAACATTTGTCAGCAATAAAAAACATCATCATG 2083
570 CCCAGTGTTCTGCTGTCGACGAGAGCGTGTCTGTTATTTTACGCGCGGCGCCACC 629
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2082 CCCTACTGCTCAATCTGACTACTGTAAATGCTCCCGCAGTTCCAGGCAAAACCCACCA 2023
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
630 ACAGCCAGCGAAATGATTTCCCTTCCATCTTTCAGCTGATACATACACGAGCATATAAT 689
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2022 ATCAACCAACAAGTGGATCCCC----- 2000
690 TCATGTCCTTTTCGGGAGGTAGCATCCCACTGACGATTAAGCGGACATTTGCTGCT 749
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1999 -----CCCTTTTTCGGGCGCAACATTTCTACTGAGCAATTTACGGGCGATCCTCTCGG 1945
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
750 GATCAGCCCGGCGGTGATATGACAGGGTAACGGTCGCGATGCGCTTCATTATSCAATGCG 809
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1944 GAGCCAGCCAGGAAGACACTCGGAGGCAAACTGTTTCGGGCAATCAACGATATAATCC 1885
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
810 GGCAGTCGAAACCGGTTGGAATAACCGTTACCGGTGCTGCTGACACTTCCGCCATCAGA 869
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1884 TGAATATCAAGCCGCGGCGCACACAGTAACCGAGTTTCGACATCTCTTCTGCTCCAGA 1825
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
870 TCGCCATCATGGTGAGATAGGCACACAATAAATACACAGATAATTTCAGGGAAC 929
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1824 TGTGCTTCATCTGTTCCGGGAAACATCACTTCATCGCAAAATGATTATTGAGAAA 1765
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
930 GTTCTGCTTACGGGTGATGATGAGTTTGTCTGACAAATAGTGAAGCGGTG---ACAG 986
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1764 AATTTTGTGTTTCGTGTCAAATATGCTTTTGCCTGATAATTCTGAATGATGTTTCCAG 1705
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
987 CATATCAGAGCGCTCAGTCTCTATATTAATGCTGATGCTGACCTATGGCAGATGACGAGA 1046
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1704 GTAAACACCGTACCAACCAACAATATTACTATCATGTCCTGTGACAAACGATTGCA 1645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1047 TCAGGTTTAAATTCGCCGATAATCCGTCGAAGTCTGAGATGGAAGAGAGTGAAGGCTG 1106
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1644 TCGGCTGAAACTATGAACTATTCGAGTAATCTCCATACAGTAGGATGTGAAGACTG 1585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1107 TTCTGAAAGGAATAAATGATGATGATGCTGCTCTTTTCTGCTTCCGAGCAATTTTA 1166
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1584 TTTGGAATAAACCGAATGTGATATCAATTTCCAAACTTACTGCTTCAAAGCAATTTTG 1525
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1167 CTATTTTCTCTGCGAG 1181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1524 CTGTTCTCCTCGAG 1510
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

SHEVIRK 1642 bp DNA linear BCT 03-FEB-1999
Shigella flexneri plasmid pMYSH6000 virulence protein (virK) gene,
complete cds.
D11025
D11025.1 GI:216943
virulence protein.
Shigella flexneri plasmid:pMYSH6000 DNA.
Shigella flexneri
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Shigella.
1 (bases 1 to 1642)
Nakatsu, N., Sasakawa, C., Okada, N., Tobe, T., Fukuda, I., Suzuki, T.,
Komatsu, K. and Yoshikawa, M.
Identification and characterization of virK, a virulence-associated
large plasmid gene essential for intercellular spreading of
Shigella flexneri
Mol. Microbiol. 6 (16), 2387-2395 (1992)
93023866
Submitted (27-APR-1992) to DDBJ by:
Nihoru Nakata
Institute of Medical Science
University of Tokyo
4-6-1 Shirokanedai, Minato-ku
Tokyo 108

```

```

Japan
Phone: 03-3443-8111x252
Fax: 03-3443-3893
Location/Qualifiers
1. .1642
/organism="Shigella flexneri"
/db_xref="taxon:623"
/plasmid="plasmid pMYSH6000"
373. .1323
/gene="virk"
373. .1323
/gene="virk"
/codon_start=1
/transl_table=11
/product="virulence protein"
/protein_id="BAA01778.1"
/db_xref="GI:216944"
/translation="MFSVSNLSIFGFLKRVFSDSLPGKWEHRKFRFMYILRCALNP
VASIRYVYELRSIQCTEDILAIQPTLPARITRPLVHLKGGRAWRSQVILHYRFVONL
PEKYSBELFPQKSVLSVQFICKGDEDFIQCSGDFDREGELMSLFFNKKIVIAITLF
SVILQNGHFAIFGGLOGAPKNTGPDILIRCATRACYGFLFPKRIIFEAFCALMKACNVS
ECLAVSEHSVIFROLRYWOKRKTFAVVSDFWESVAGTKCGDWYKLPQTQVVRKPLSN
IASKKRSEYRKRYALDLYIHETAIRSLDAPVNSEHQDLN"
BASE COUNT 451 a 311 c 376 g 504 t
ORIGIN

Query Match 13.5%; Score 159.4; DB 1; Length 1642;
Best Local Similarity 65.1%; Pred. No. 6.7e-40;
Matches 235; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 210 TGACTAGCCAGACACAATGATCTGTCCGCTTCTGTTAATATCAAAACCGGTACTCAATA 269
D 210 TGACTAGCCAGACACAATGATCTGTCCGCTTCTGTTAATATCAAAACCGGTACTCAATA 269
Db 361 TGCTTCGCTAATGTAGTATTTTAAACGAGTATTTAATATCAAAACCGGTACTCAATA 302
QY 270 TCTTCTCGCTGGCTGGCTGCATCATCCGGAAGGCTTCCGGTGGGATAAAATACCGC 329
D 270 TCTTCTCGCTGGCTGGCTGCATCATCCGGAAGGCTTCCGGTGGGATAAAATACCGC 329
Db 301 TCCTGTTTCCCTGGCGAGCCATCTGATAAAACGCCAGGCTATTAAAAAATCATTC 242
QY 330 AGTGGCGCGGTCCATGACAGACATCCGCCAGGGTAACAGCGTCCCTGTACATTTCTC 389
D 330 AGTGGCGCGGTCCATGACAGACATCCGCCAGGGTAACAGCGTCCCTGTACATTTCTC 389
Db 241 AGGCGGCACATCCATGCTGCTTATTACCTGCTGTAAACATGTCCCGGTCTGGTTATTC 182
QY 390 TGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGCGACCGCGGAGACTGAC 449
D 390 TGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGCGACCGCGGAGACTGAC 449
Db 181 TGAATACATCAGGATCCCAACATCTGACTGCGGAGTACAGGACAGAGAAATGCCGAT 122
QY 450 GCTTCAGCAGTACATACCAAGCTTCATTTCCGAAGGCTATGACACACACTGGCA 509
D 450 GCTTCAGCAGTACATACCAAGCTTCATTTCCGAAGGCTATGACACACACTGGCA 509
Db 121 GCTTCGCCAGTACATACCAAGCTTCATTTCCGAAGGCTATGACACACACTGGCA 62
QY 510 ATCCGGTAGCCGTTACGCTGGGAAAGGCGACCTGCCATTAAACATCTCCGCTCAT 569
D 510 ATCCGGTAGCCGTTACGCTGGGAAAGGCGACCTGCCATTAAACATCTCCGCTCAT 569
Db 61 ACCGATATACGGGGCGGAGGAAACATTTGTCAGCAATAAAACATCATCATGATG 2
QY 570 C 570
D 1 C 1

RESULT 14
E06635/c
LOCUS E06635 Shigella virK gene. 1642 bp DNA linear PAT 29-SEP-1997
DEFINITION Shigella virK gene.
ACCESSION E06635
VERSION E06635.1 GI:2174818
KEYWORDS JP 1994030766-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1642)
AUTHORS Yoshikawa,M.
TITLE ATTENUATED LIVE VACCINE OF DYSENTERY BACILLUS HAVING HIGH SAFETY
JOURNAL Patent: JP 1994030766-A 1 08-FEB-1994;

RES DEV CORP OF JAPAN
OS Shigella flexneri 2a
PN JP 1994030766-A/1
PD 08-FEB-1994
PF 15-JUL-1992 JP 1992210772
PI YOSHIKAWA MASANOSUKE
PC C12N1/21,A61K39/00,A61K39/00,A61K39/112,A61K39/112,C12N15/31,
C12P21/02,
PC (C12N1/21,C12R1:01),(C12P21/02,C12R1:01);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=YSH6000T;
CC *source: clone=pMYSH6000;
FH Key Location/Qualifiers
FT mat_peptide 373. .1320 /product="Shigella virK protein".
FT 1. .1642 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 451 a 311 c 376 g 504 t
ORIGIN

Query Match 13.5%; Score 159.4; DB 6; Length 1642;
Best Local Similarity 65.1%; Pred. No. 6.7e-40;
Matches 235; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 210 TGACTAGCCAGACACAATGATCTGTCCGCTTCTGTTAATATCAAAACCGGTACTCAATA 269
D 210 TGACTAGCCAGACACAATGATCTGTCCGCTTCTGTTAATATCAAAACCGGTACTCAATA 269
Db 361 TGCTTCGCTAATGTAGTATTTTAAACGAGTATTTAATATCAAAACCGGTACTCAATA 302
QY 270 TCTTCTCGCTGGCTGGCTGCATCATCCGGAAGGCTTCCGGTGGGATAAAATACCGC 329
D 270 TCTTCTCGCTGGCTGGCTGCATCATCCGGAAGGCTTCCGGTGGGATAAAATACCGC 329
Db 301 TCCTGTTTCCCTGGCGAGCCATCTGATAAAACGCCAGGCTATTAAAAAATCATTC 242
QY 330 AGTGGCGCGGTCCATGACAGACATCCGCCAGGGTAACAGCGTCCCTGTACATTTCTC 389
D 330 AGTGGCGCGGTCCATGACAGACATCCGCCAGGGTAACAGCGTCCCTGTACATTTCTC 389
Db 241 AGGCGGCACATCCATGCTGCTTATTACCTGCTGTAAACATGTCCCGGTCTGGTTATTC 182
QY 390 TGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGCGACCGCGGAGACTGAC 449
D 390 TGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGCGACCGCGGAGACTGAC 449
Db 181 TGAATACATCAGGATCCCAACATCTGACTGCGGAGTACAGGACAGAGAAATGCCGAT 122
QY 450 GCTTCAGCAGTACATACCAAGCTTCATTTCCGAAGGCTATGACACACACTGGCA 509
D 450 GCTTCAGCAGTACATACCAAGCTTCATTTCCGAAGGCTATGACACACACTGGCA 509
Db 121 GCTTCGCCAGTACATACCAAGCTTCATTTCCGAAGGCTATGACACACACTGGCA 62
QY 510 ATCCGGTAGCCGTTACGCTGGGAAAGGCGACCTGCCATTAAACATCTCCGCTCAT 569
D 510 ATCCGGTAGCCGTTACGCTGGGAAAGGCGACCTGCCATTAAACATCTCCGCTCAT 569
Db 61 ACCGATATACGGGGCGGAGGAAACATTTGTCAGCAATAAAACATCATCATGATG 2
QY 570 C 570
D 1 C 1

RESULT 15
AF326777
LOCUS AF326777 Shigella flexneri 2a 66714 bp DNA linear BCT 18-JUN-2002
DEFINITION Shigella flexneri 2a SRL pathogenicity island, complete sequence.
ACCESSION AF326777
VERSION AF326777.2 GI:21450881
KEYWORDS
SOURCE Shigella flexneri 2a.
ORGANISM Shigella flexneri 2a.
REFERENCE 1 (bases 1 to 66714)
AUTHORS Luck,S.N., Turner,S.A., Rajakumar,K., Sakellaris,H. and Adler,B.
TITLE Ferric dicitrate transport system (Fec) of Shigella flexneri 2a

```

YH6000 is encoded on a novel pathogenicity island carrying multiple antibiotic resistance genes
 Infect. Immun. 69 (10), 6012-6021 (2001)
 21437601
 11553538
 2 (bases 1 to 66714)
 Luck, S.N., Turner, S.A. and Rajakumar, K.
 Direct Submission
 Submitted (06-DEC-2000) Microbiology, Monash University, VIC 3800, Australia
 3 (bases 1 to 66714)
 Luck, S.N., Turner, S.A. and Rajakumar, K.
 Direct Submission
 Submitted (17-MAY-2002) Microbiology, Monash University, VIC 3800, Australia
 Amino acid sequence updated by submitter
 4 (bases 1 to 66714)
 Luck, S.N., Turner, S.A. and Rajakumar, K.
 Direct Submission
 Submitted (18-JUN-2002) Microbiology, Monash University, VIC 3800, Australia
 Sequence update by submitter
 On Jun 18, 2002 this sequence version replaced gi:15808696.
 Location/Qualifiers
 1. .66714
 /organism="Shigella flexneri 2a"
 /db_xref="taxon:42897"
 /note="corresponds to Escherichia coli min 23"
 458. .66714
 /note="SRL pathogenicity island"
 619. .1836
 /gene="int"
 619. .1836
 /gene="int"
 /note="member of the P4 bacteriophage integrase family"
 /codon_start=1
 /transl_table=11
 /product="integrase"
 /protein_id="AA08426.1"
 /db_xref="GI:15808697"
 /translation="MAVLTDTKARIHKDPDKPLPHGGITGLTLHPSSVKGKRGKWFYRVSPVQKRRNAGLTVPEVSIAGEARTIMRQOLAAGDDPDLKKAESKVVIPPTFA DARRVRAELPGWENPKHVOMLSLENVAPQLGAKTLDSTADVAETLRPWLT LSETASVKGRIHVQWQWAGFCVANPDVVDVHLLPQTRGRDEHPAMPWRQLPL FVAISVSDPEYNTRALLMWILTATRSGEARGMAEIDFHKRWVTIPAEKMKARL QHVPVLSRQAIYILENRGLHDELVPSPKQQLSDVLTSLFKKKAVSDIPGRVA TAHGFRSTFDWQSEQYSRDLAERLAHTLKNKEAAYHRTDLLEQRPVMMQAWADY VMSQAKNTKEFA"
 complement(2006. .3535)
 /note="Orf2; similar to yfji on the cryptic P4 prophage
 CP4-57 of Escherichia coli K-12"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AA08427.1"
 /db_xref="GI:15808698"
 /translation="MMTKLSLPDLTKLPDRQALLDEMHEETGIRREILLVMLT VMAASVDTRHEVLSGGQRTSLQIFWCLSSAGSGKTSACAKLIAPVHTEELHOAY IDDKNNDRMWMTTQDKTLERRYKEIERSPENAAARALEKCIANKPEPPVQOV LIYNDATPEGIALKSGSPSLLSDEGGTILDKRFRKSAITLWSGQPVNVERAS RPTGRVNDVRLTILITQPVIFNKFLLTGDDQIRNGFLARLLFCPEGPAKIMTTEPH TATPVVQQACASRTKFGSGQISRLASRRRAKSKERICMTLSNAASRALERPHK EMNDTQNMHMTGFEDIIVKREQVVRILAELEKDPDSTVITLESNSAIYLLEF YFKHLIYKLSREISPAEKLDKWLQRIITTAGYIFOKSYIIQVAPYALRKCKVLDE ALDILAQRKIRIDDLNVVYIGNITPSELAKKLINIPAFDAGVFCIDHQNLKLYHNR L"
 complement(3507. .3812)
 /note="Orf3; similar to AlpA on the cryptic P4 prophage
 CP4-57 of Escherichia coli K-12"
 /codon_start=1
 /transl_table=11
 /product="putative transcriptional regulator"

/protein_id="AA08428.1"
 /db_xref="GI:15808699"
 /translation="MHMNNMNSNOTGQYTPSRILRMAELRLLGISRSTIYVKINPA SKYHPTPKPIRLGSAVGVGRASAIDEWMLHTAPAWSEPNKELNDDDKITP"
 3868. .5730
 /note="Orf4"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AA08429.1"
 /db_xref="GI:15808700"
 /translation="MOTNRQAQMFTFMNLYDICTVLGPKGENPAHKRKSNTFEM HQKWDQRYNEISRIIDAEGVFSQQRRIIYARVEHYFYMNSPVHSTLKPFLRSY CLRTGVLFLVVDYNTIPENDSAFYHIYNFLQKSYCPCLDHADTESDAAYKRLV REYLAELGENKDFHENGKALGKRYGTGIRKDNKSKSLMQQYIMAIKNVKKDYRE KKLDELKFNLDKFNNAFYSLSVLDDIQRTKILQSLAYLVIRGLWHGL YGAAQYLYDFTSFDTPIYAKLLEMFYQFONSAEGTSLRSYSLDDSKQYIISRLKD LVFNINDKNGCDDAYLKLIISYFGOLQNEAVHVTSCYETLAVYICLIHKKNINDVLOH YDDMRKGLFGELPSGVYRGALSLLRTALEVKNRKNIKYSLFYWLHDVAVQDAFI EKIPLDIPYKEGEIQYDANNFTLMRVIKMNCLEKISTKPYITAPPYITGLDDVEK VLDKINILDKREYVDGKTLAEVIMENKVLSSRERKETMIGLTGSKKTYTLQCVKEL GVLVHYKSPVDEIKNVMLYGDKAENRRRMIVDALTIICEDDIRNPPFLS"
 7104. .7697
 /note="Orf5"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AA08430.1"
 /db_xref="GI:15808701"
 /translation="MYNANPNYKMDFMILKDVNEHMDGMFORFKLLPFRIDFAYRKD TPFSGHCHSHCMCEIYRLSETQTLAGYVWMEYTPDKGLHIFHIGYLDGQRHKN YQISROLQDWIRITEGEGYFHLCKRAKDKYVPRIDVHYSDKSAVDLRLVALSYLAK ODKHEGIIIGRSRLPEKSNRGRPDITELWGLMSELS"
 8449. .8745
 /note="Orf6; similar to yfjJ on the cryptic P4 prophage
 CP4-57 of Escherichia coli K-12"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AA08431.1"
 /db_xref="GI:15808702"
 /translation="MSLSRLPFFNKDAYHYHGNYSRDTLRLGLITGAWYSALGLELED HGVLVHFANGRYILDINDPNFPEQYQRLVRLDYTLTKDTKASQGRNFGYSGF"
 complement(9112. .10421)
 /note="IS629-like element"
 10422. .>10905
 /note="remnant of IS200-like element"
 complement(10980. .11918)
 /note="Orf7; similar to a LysR-like transcriptional regulator"
 /codon_start=1
 /transl_table=11
 /product="putative transcriptional regulator"
 /protein_id="AA08432.2"
 /db_xref="GI:21450882"
 /translation="MKNITKWLXDFTLEACRHFSAAKERNLSQAFSRRIRKALEA AIGVLPDRTTPTLQTEGKLFHSQTSRLQLQECNLGELGSLGLGVPNIKTAHAH SLSLVPRLVHSLTAYGGEFYHVEADIVQAVNTLRGSKDFIISFRDRLMQSPF CCLIFESLPEVPCAAQAQHPFDITQPVPLNTATSYMGLVNRHLAEVGGITG RTIFTSMSSELLKNMALNGYIATWPIASIVDELQTKRLICLDAAKLTVPVQAYIRM NTRNRTAENWRILQEHMPDDLIQOISMEPPARN"
 12299. .12988
 /note="Orf8"
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="unknown"
 /protein_id="AA08433.1"
 /db_xref="GI:15808704"
 /translation="MGPATVDLPNKFVTFVTAACQDQEHPLIISISIPDIPDRTDALM

CDS	<p>HGSRPLPVWRDYMVKHLEDAGRCIVIPONTAHWKELKDACHIDILSIVETTINEV RACKTRIGGLATNATIMGLYOGKIESLGTCVSPDADGQKKVMYSIYSLAGNATIA AKMLNPGAEITLFSRGAEIILVLCGTEIPVILAQAVREQPSRYIDTSTASLVRAKIKWYE NRVGKHYLLTQ"</p> <p>13054..14352</p> <p>/note="orf9; similar to DcuA in <i>Escherichia coli</i> K-12"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="putative anaerobic decarboxylate transporter"</p> <p>/protein_id="AAL08434.1"</p> <p>/db_xref="gi:15808705"</p> <p>/translation="MVMLEITIVVLGAIFGIROGGIGIGLGGGLAILTLGFLGPMG SPVDVILITMVTVAASALQAGGMDYLRVAGNMRPNPKYINIIPIITWSMTIM AGTCFIVFSTLPVTAEVAKESGIRPSTLAGSVASQALSGSPISAAAMLTMEV NGSVFIQVMSVCLPTSFARAAVAFATSRGCELDQDDVEYLERLQGLVKQKVENNSSI KPGATLSVGLFLLATVAIVLAIAFFQRPGRFDIGKPMETRDIIICMLSGAACLMMVLC KATDAIILSTFRAGMSLAIVLQIGVITGTFIDAHLEIKDINGADIIQAOVLMMLAL VLFGTCALLYSQGAFTNIIPLAVALVDPTWAILASVAVTGVFVLPTPTTSLAAEF DTTCTTRVGRVLYNHPMLPGLGGIITAGVALGEVIAPMII"</p> <p>complement(<14553..>15322)</p> <p>/note="remnant of IS600-like element; central region deleted"</p>
repeat_region	<p>complement(15323..16090)</p>
repeat_region	<p>/note="IS1 element"</p>
gene	<p>complement(16171..16962)</p>
gene	<p>/gene="aadA1"</p>
CDS	<p>complement(16171..16962)</p> <p>/gene="aadA1"</p> <p>/note="confers streptomycin and spectinomycin resistance; similar to AadA1 from Tn21 on plasmid R100"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="streptomycin adenylyltransferase AadA1"</p> <p>/protein_id="AAL08435.1"</p> <p>/db_xref="gi:15808706"</p> <p>/translation="MRSVIAEVSSTOLSEVVGVIERHLEPTLLAVHLYGSADVGGILKP HSDIDLLVTVTVRLDEFTTRALINDLETSSAPGESEILRAVEVTIVHDDIIPWRYP AKRELQGEKWRNDILACIFEPATIDIDLLAKTRAEHSLAVETVGPAAEELDPVPEQ DYDFXLNETTLWNSPDDKGRNXYVTLXIXWSKXTKAPKDDXADWAMERLPA QYQVLEHROXYLQSPEDRAVQDLSEFVHYVKGEITKVVVG"</p> <p>complement(17075..17905)</p> <p>/gene="oxa-1"</p>
gene	<p>/gene="oxa-1"</p>

Query Match	13.1%	Score 154.8	DB 1	Length 66714	
Best Local Similarity	58.2%	Pred. No. 4.8e-38			
Matches	292	Conservative 0	Mismatches 207	Indels 3	Gaps 1
QY	683	CATAAATTCATGCTCTTTTTCGGGACGTAGCATCCCACTTCAACGATAACGGGAACATT	742		
Db	31985	CTTAAATTCGTGCCCTTTTTCGGGCGCAACATCTTACCTGAGCAATTAACGGGGCATCC	32044		
QY	743	GTCTGCTGTGATGACGCGCGTGTGATATGCAAGGGTAACGGTGCATGGCTTCATTATG	802		
Db	32045	TCTTCGGGACGCCAGCCAGGAAGAACAACACTCGGAGGCAAACTGTTCCGGGAATCAACGTA	32104		
QY	803	CAATCGGGCCAGTGGAAACCCGGTGGATTAACCGTTTACCGGTGCTCTGACACCTTCGCG	862		
Db	32105	TAATTCCTGGAAATCAAGCCAGGAGGCACACAGTAACCCGAGTTCGACATCCTTCTGTG	32164		
QY	863	CATCAGATCGGCCATCATGGGTGAGATAGGCACACAATGAAATCACACAGATAAATTCAG	922		
Db	32165	CTCCAGATGTCTTCATATTGTTTCGGGGAACAATCACTCATCGCAAAATGATTTAT	32224		
QY	923	GGAAAAAGTCTCTGGTCTTACGGGTGATGTAGGTGTTTTGTCTGCAATAGTGAACGGGTG	982		
Db	32225	TGAAAAAATTTGTTTTTCGTGTCAAATATGCTCTTTTGCCTGATAATTTCTGAATGGATG	32384		
QY	983	---ACAGCATATCAGACGGCTCAGTCTCTGTATATTACTGTTCATGGCCACTATGGCAGAT	10399		
Db	32285	TTTCCAGGTAAAAACCGGTACCAACCCACAATATTACTATCATGTCCTCCTGTGACAAAC	32344		
QY	1040	GACCAGATCAGGTTTAAATTCCTCCCATTAATCCGTCCAAAGTCTGAGGATGGGAAGGATG	1099		

Db	32345	GATTGCATCGGGCTGAAAACCTATGGAACCTATTCGAGTAACTCCATACAGTAGGAATGTG	32404
Qy	1100	AAGGCTGTTCTCGAAAGGAAATAAAGTGCATCATGCCCTCTTTTCTGGCTTCGGAGC	1159
Db	32405	AAGACTGTTTCGAAATAAGCGCAATGATCAATTCCAACTTACTGGCTTCAAAGC	32464
Qy	1160	AATTTTACTTTTTTCTCTGCAG	1181
Db	32465	AATTTTGCCTGTTCTCCCTGCAG	32486

RESULT 16	BACJOJC/c	LOCUS	BACJOJC	6540 bp	DNA	linear	BCT 26-MAR-1996
DEFINITION	Bacillus subtilis dihydropicolinate reductase (jojF) gene, complete cds; poly(A) polymerase (jojI) gene, complete cds; biotin acetyl-CoA-carboxylase ligase (bira) gene, complete cds; jojC, . . . jojD, jojF, jojG, jojH genes, complete cds's.						
ACCESSION	L38424						
VERSION	L38424.1 GI:755600						
KEYWORDS	biotin [acetyl-CoA carboxylase] ligase; dihydropicolinate reductase; poly(A) polymerase.						
SOURCE	Bacillus subtilis (sub_strain 168, strain Marburg) DNA.						
ORGANISM	Bacillus subtilis						
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.						
AUTHORS	Bower,S., Perkins,J., Yocum,R.R., Serror,P., Sorokin,A., Rahaim,P., Howitt,C.L., Prasad,N., Ehrlich,S.D. and Pero,J.						
TITLE	Cloning and characterization of the bacillus subtilis bira gene encoding a repressor of the biotin operon						
JOURNAL	J. Bacteriol. 177 (9), 2572-2575 (1995)						
MEDLINE	95247697						
PUBMED	7730294						
REFERENCE	2 (bases 1 to 6540)						
AUTHORS	Sorokin,A.V., Azevedo,V., Zumstein,E., Galleron,N., Ehrlich,S.D. and Serror,P.						
TITLE	Bacillus subtilis chromosome between spoIIA and kdg loci .						
JOURNAL	Unpublished (1995)						
FEATURES	Location/Qualifiers						
source	1..6540						

```

terminator
  /map="202 degrees"
  29.42
/standard_name="transcription terminator"
/note="putative"
complement(59..706)
/gene="jojC"
complement(59..706)
/gene="jojC"
/note="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92872.1"
/db_xref="GI:755601"
/translation="MFYVTLGVTFVSLSLFISFQRYEIHMP
ALGLIIFQRTGGVDGIIARLVNKKYFGIPMGRTMFAFDK
TLVAVFVAARLDIIQIGGGYAAKGATTISSKNDLIQKKILE
EDIDLVCVVPKKNELVMLKLSVINSIDPHAFVAVSDVHDVLG
1011..1346
/gene="jojD"
1011..1346
/gene="jojD"
/note="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92873.1"
/db_xref="GI:755602"
/translation="MSDKTWKDIQAEVDVRYIGQKGYGSP

```



```

VNHRYGKPKKATEDDKSMEEIGDVLFLVCLANSIDISLEEAHDRVHMKFNTRDKD
RWRKEEGK
1358. .2101
/gene="dapB"
1358. .2101
/gene="dapB"
/function="lysine biosynthesis"
/notes="28.2% of identical aminoacids with the Escherichia
coli dihydrodipicolinate reductase; putative"
/codon_start=1
/transl_table=11
/product="dihydrodipicolinate reductase"
/protein_id="AAA92874.1"
/db_xref="GI:755603"
/transl_table="MSNETIKLVIAGPRGRMGQAVKLAERTPHFDLVGAIDHTYDQ
KLSDVVESDAFYTDIHCFTPTQDVLIDLTPEIGKVHTKIALEHGVVRPVGTT
GFSEADLKLTSTTEEGIGAIIPNFAIGALIMKFKSMAANYFEDVEIIEHDDK
LDAPSCWALKAEMI SVRKEKQGHDEKELLPGARGAQNGIRLRLHRIQRHDSYNR
ASPMGKLSVEQVMKIDQLVYLENIID"
2117. .2530
/gene="joJF"
2117. .2530
/gene="joJF"
/notes="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92875.1"
/db_xref="GI:755604"
/transl_table="MKIALIAHDKKKQDMVQFTYAVRDILKNHDLVATCTTGKIHFA
TGQIEPFOGSLGSDQOIGALIAANALDLVFLRDLPLAQHPEDVSALIRLDVYS
IPLATNNGTRAILVRLDEGVFEFRLLRGEENV"
2523. .3197
/gene="joJG"
2523. .3197
/gene="joJG"
/notes="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92876.1"
/db_xref="GI:755605"
/transl_table="MYNADVLAFGASHDDVEIGMGTTIAKFKQKVKMVICDLTAEAL
SSNGTVSLRKEEAARILGADKRILQILPDRGLMSDAQRSIVTVTRICRPAVF
MPYKDRPHGNAALYVEAIPFSGIHAKDEKSLPAHKVSKVIYIMINGPHQDFV
IDISDTIARKQSLNAYKQFIPSKDSVSTPLTNGYIETVEAREKLYGREAGVEYAEV
SFPNGC"
3229. .4302
/gene="joJH"
3229. .4302
/gene="joJH"
/notes="21.6% of identity to trans-acting transcription
factor of Saccharomyces cerevisiae; 25% identity to sucrose
synthase of Zea mays; putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92877.1"
/db_xref="GI:755606"
/transl_table="MRKLGKLAEGHEIHFTSSIPRLNTHPHNIHFEVENVQYA
VFYPPYDITLASKIAEVAERENLDLIHAHVLPRAVCAYLAKOMLKRNGIVTTVLHG
TDITVLGYDPSLDLIRFALESDDRTYAVSSALAAETDILKPEKTIETIYNFIDERV
YLAKNAAIKERHGILPDEKVVHVSFRKRVQDVIRVFRNIAGKAKALLVGDG
PEKSTACELIRKGLQDQVLMGNQDRVEDLISDLKULLSEKESFGLVLEAWACG
VPCIGNIGGIPVIVKNVSGELVDVGDVTAATARAMSILEDEQLSNRFTKAAIEMLE
NEFSRKKIVSQYBQIADLAEP"
4307. .5440
/gene="paps"
4307. .5440
/gene="paps"
/notes="38.9% of identical aminoacids with the Escherichia
coli pcnB gene product, encoding poly(A) polymerase;
putative"

```

```

/codon_start=1
/transl_table=11
/product="polyA polymerase"
/protein_id="AAA92878.1"
/db_xref="GI:755607"
/transl_table="MEKVFITKALPVLIRILBAGHOAYFVGAVRDSYMKRTTIGDVIDIA
TDAAPQDVERLFTQVDYDKDSYDVPFRSEVQFISSLEEDLKRRLDTINAMAMTA
DKGLVDYFGGKKIDQKIVTVGKPEDFQEDALRMLRAVRFMQLGFTLSPTEEEAI
AREKSLLSHVSVERKTIEFKLQGRASQALOTLIOTRLYEELPGVHKRENLISTS
EPFPLSTSRLELNAALINIGLVLKADAPLFLKAWKLPKVIKEAHIADTFGQSLDA
WMYRAGKALLSAKISQLRQNEKLDKLDIQYATQNLPIKSLDKDLDTIGKDLILA
LNNRPAKWSELEQWIEQAVVTGKLSNQKHIEWLKTCQGH"
5425. .6342
/gene="bira"
5425. .6342
/gene="bira"
/function="bifunctional biotin operon repressor"
/notes="26.7% identity to the Escherichia coli bifunctional
biotin operon repressor"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="biotin [acetyl-CoA carboxylase] ligase"
/protein_id="AAA92879.1"
/db_xref="GI:755608"
/transl_table="MRSTLRKDLIELFSQAGNEFTSGOKISDALGCSRTAVWKHIEEL
RKEGYEVAERKRGVRLKPKGLSEIRFGLKTEVMGOHLIYHDLVSSQTOKTAHEL
ANNNAPEGTLEGNGVWMSLILRPDIPLOKTPQLTLLAAVAVVQIEEAGTQTDKWP
NDILGKKTGVLTEMOAEDRVSRTIGIGINVNQPNDFPDELDIATISLSQAG
EKIDRAGVQHILLCFERKRYDMTHGFTPKLLWESLALGIGTNMRARTLNGFTYFK
ALGIDDEGVLLLENEGKIKIYSADIELG"
6501. .6533
/notes="putative"
termInator
BASE COUNT 2036 a 1318 c 1547 g 1639 t
ORIGIN
Query Match 5.4%; Score 64.2; DB 1; Length 6540;
Best Local Similarity 54.4%; Pred. No. 9.9e-09;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 335 GCGGTCTCATGCAGACATCCCCACGGGTAAACAGCGTCCCTGTCACATCTTCTGAAT 394
Db 4158 GGCTTTGCGGCGGTAAACATCACCAGCATCCACAAAATCCGCTCACATTTTAT 4099
QY 395 GACATCAGGATCCCGCGTCTCACTGGCGATACGGCGACGGCGGAGACTGACGCTTC 454
Db 4098 AACCTCAGGATACCGCAATGTTTGTTCATAACAGGCACCTCGCAAGCATCGCTTC 4039
QY 455 AGCAGTACCATACCAACGGTTTCATTTCCGAAGGCATGACCACACATGGCAATCCG 514
Db 4038 AAGCAGGACAAAGGCCAAGCTTTCTTTTCAGATAGCAGCAGCTTCAATCGCTAATAGA 3979
QY 515 GTAGACCGGTAACTGCGGAAAGGCGACCTGCCATTAACACATCTCGGCTCATCC 571
Db 3978 ATAAAGATCTTCAACACCGTCTTGATTTCCAGCATTAAGACTTGGTCTTCCAAGCC 3922
RESULT 17
AX100832
LOCUS AX100832 10801 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 76 from Patent WO0121772.
ACCESSION AX100832
VERSION AX100832.1 GI:13619763
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 10801)
AUTHORS Yocum,R.R., Patterson,T.A., Hermann,T. and Pero,J.G.
TITLE Methods and microorganisms for production of panto-compounds
JOURNAL Patent: WO 0121772-A 76 29-MAR-2001;
OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES Location/Qualifiers

```


[illegible]

```

3379. 4053
/gene="qcrB"
3379. 4053
/gene="qcrB"
/function="cytochrome reductase"
/standard_name="cytochrome b6"
/note="42.3% of identical aminoacids with Chlorella
prothecoides chloroplast cytochrome b6"
/citation=[2]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="cytochrome b"
/protein_id="AAB38436.1"
/db_xref="GI:1146228"
/translation="MLNKKIYDWDERLIDITPMRDIDADHEVPEHVPNPHFSFYTCF
GGLFFVTYVIGSMFLTYVPDIKNWSEVYIQNEVAFGQIVRGMHHWGASLVI
VMFLHTRVFFOGAYKKPRELAWIVGLVYFMGLGFTGYLLPMDMAKALFATKVL
QIABATPLIGTQVKTLAGHDIVGAOTLRFPAIHVFFLPAALFGLMAHFIMIRKO
GISGPL"
4088. 4855
/gene="qcrC"
4088. 4855
/gene="qcrC"
/note="29.1% of identical aminoacids with the Spinacia
oleracea chloroplast cytochrome b"
/citation=[2]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="cytochrome b"
/protein_id="AAB38437.1"
/db_xref="GI:1146229"
/translation="MHRGKGMKFGVDSRIPAEKKPNPKDYSEVPGKTEAFWPNELK
EMYGAVFLGFLVITVHPPLERMDPTDTGVIPLPDWYFLFLQLLXEYAAASF
TVGAMIPGLAFGALLAFDLDRGTERPWRKPVAVGMMLLALSAAVLTWGSVATH
DMAKAEQKTKPADIDTNAEGYKVFKEQCISCHGNDIQQGAAGPSLVDGLKPD
IKIAVEGKMPAGVFGKNDKQLELAKFISETTAK"
4898. 4922
/standard_name="transcription terminator"
/note="putative"
/citation=[2]
5002. 5559
/gene="ypjA"
5002. 5559
/gene="ypjA"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/db_xref="GI:1146230"
/translation="MLILVAINFLGVGYWYVLPOLLETPARFLIFVDSPTATFF
FLVQLLAFMRNAPLLEALVTLVKYGLWAVAMNFIPLVATGDLPEWGYMLIASHF
AMAVQGLVSPYFSPFHWHLAIAAVTLHNDVIDYLFDMMPQVYSLDSLTETGTYGF
WLSIFSIALAYFLVSKKQTKLELM"
5628. 6422
/gene="ypjB"
5628. 6422
/gene="ypjB"
/function="hypothetical"
/note="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/db_xref="GI:1146231"
/translation="MKRLITCLLIALIFYNGNAAGRSLEELNLSDTVFQMTFRQ
AKYEFALQVLEFFKTLKSAEKQODPMLTGAQITQITLGYNDMVRSLKQADSDTK
LRAAQFRMLADVNDSDPLWGLSKRPIMEAFTELKRDVKNGSTSFHEKWNFISL
LVIPLPSITIDVSEQLETGVKHIDVIEQEFQMTSTKLERLSLQHLKDNVDRV
EDDDADPSLLWITTSIIITLTYVYVIRKAKKKNLKKRDYK"

```

```

terminator
6434. 6447
/standard_name="transcription terminator"
/note="putative"
/citation=[2]
complement(6464..7111)
/gene="ypjC"
complement(6464..7111)
/gene="ypjC"
/function="hypothetical"
/note="putative"
/citation=[2]

Query Match 5.4%; Score 64.2; DB 1; Length 23775;
Best Local Similarity 54.4%; Pred. No. 1.3e-08;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 335 GCGGTCATGACGACACATCCCGACGGGTACAGCGTCCCTGTGCACATCTTCTTGAT 394
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10743 GGTGTGTGGCGCGGTACATCACACATCCCGACATCCCAAAATCCGCTCACATTTT 10684
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 395 GACATCAGGATCCCGCGCTCTCACTGGCGATACGGGCACGCGGACGTGACGCTTC 454
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10683 AACCTCAGGATACCGCCCAATGTTTTCATACAGGACATCCCGACGATCGCTTC 10624
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 455 AGCCATACCATACCAACGCTTCAATTTCCGAAGCATGACACACACTGGCAATCG 514
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10623 AAGCAGGACAAGGCCAAAGCTTTCTTTTTCAGATAGCAGCGCTTCAATTCGTA 10564
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 515 GTAGCCGTACGCTGGGNAAGGCGACCTGCCATTACACATCTCCGCTCATGCC 571
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10563 ATAAAGATCTTCAACACGCTGTGATTTCCAGCATTAAGACTTGTCTTCCAAGCC 10507
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 19

BSUB0012

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BSUB0012 213680 bp DNA linear BCT 26-NOV-1997
 Bacillus subtilis complete genome (section 12 of 21): from 2195541 to 2409220.
 299115.1 GI:2634478
 Bacillus subtilis.
 Bacillus subtilis.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Conner, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Gallaron, N., Gim, S.Y., Glaser, P., Goffeau, A., Gollightly, E.J., Grandi, G., Hilsepp, G., Guy, B.J., Haga, K., Haech, J., Harwood, C.R., Henaut, A., Hildebrand, H., Holtsappel, S., Hosono, S., Hulio, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Masuda, S., Mael, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Park, V., Pohl, T.M., Portet, D., Porwollik, S., Prescott, A.M., Presecan, E., Puig, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serrero, P., Shin, B.S., Soldo, B., Sorokin, A., Taconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenecker, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,

Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and Danchin, A.
 The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*
 Nature 390 (6657), 249-256 (1997)
 98044033
 PUBMED 9384377
 REFERENCE 2 (bases 1 to 213680)
 Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
 Direct Submission
 Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur, Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr.
 adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

FEATURES
 source
 1. .213680
 /organism="Bacillus subtilis"
 /strain="168"
 /db_xref="taxon:1423"
 complement(90. .782)
 /gene="yoqL"
 complement(90. .782)
 /gene="yoqL"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13977.1"
 /db_xref="GI:2634479"
 /db_xref="SPTREMBL:O34581"
 /translation="MDSKKHKIINGYVYKNCISCKSWLPATEENFYVKKNKDGLHSY
 CKACVLKAKESMLKNDYQLEMRERNLLPGMKEAKKYNSLKKKKQOIWQERNK
 LKUNTRLRQDAHKHNTDVOVKCKDYFNKCSYGLKIEDHKILFRGTIYIOSDFH
 KEVHDKGANDISNCIPACKSCNSKHDFAFEWYNSNNKNSERLLKIKELWLNRFK
 EERQDSEWRTKG"
 complement(822. .1025)
 /gene="yoqK"
 complement(822. .1025)
 /gene="yoqK"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13978.1"
 /db_xref="GI:2634480"
 /db_xref="SPTREMBL:O34326"
 /translation="MKLIERTYEERLYIYSDDEARKHFVSELKPKGMVDCNCGVI
 DLNQIKQYFRKRELDDYVFN"
 complement(1045. .1560)
 /gene="yoqJ"
 complement(1045. .1560)
 /gene="yoqJ"
 /function="unknown"
 /note="similar to hypothetical proteins from B. subtilis"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13979.1"
 /db_xref="GI:2634481"
 /db_xref="SPTREMBL:O34359"
 /translation="MKNPTMLKLDKLLAVIEELITKENKYRITGGALGTDQAACWC
 VHLKKKHPHIKNIATPPKEDKVSADQKMYKRLNVADEIINVELDKYKVSSE
 KPGFSPAKMKRENYIDHSEIAIVAYDGSKSGTRNCLNAYAKTYLGHQLWRLHPD
 NFEDITIFVG"
 complement(1771. .1965)
 /gene="yoqI"
 complement(1771. .1965)
 /gene="yoqI"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13980.1"
 /db_xref="GI:2634482"
 /db_xref="SPTREMBL:O34339"
 /translation="MKSIIIPENTKITVGVKPINDESNTKVIAIYDREPEAFWV

gene
 CDS
 complement(2014. .2466)
 /gene="yoqH"
 complement(2014. .2466)
 /gene="yoqH"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13981.1"
 /db_xref="GI:2634483"
 /db_xref="SPTREMBL:O34499"
 /translation="MKRFILVLSFLSIIVAYPIOTNASPMPCSVILLEPVDKLNKNAK
 VALIYKVLNPPSAARTNISILAVHLPAPSSFGNYSEGFATKPGEISWRFLDIPTP
 EESPWAGRIDTISAEMKNVKQVRLSNSSQRLGPSILITKNIESCY"
 complement(2549. .2806)
 /gene="yoqG"
 complement(2549. .2806)
 /gene="yoqG"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13982.1"
 /db_xref="GI:2634484"
 /db_xref="SPTREMBL:O35030"
 /translation="MEKNLKEITFDLKNKIRPGTKVQVTYKGTETMETHNGEVII
 NNGEKFYNYVDREGYVGHCHVALDLKNYPDSLIVEIKSK"
 complement(2851. .3054)
 /gene="yoqF"
 complement(2851. .3054)
 /gene="yoqF"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13983.1"
 /db_xref="GI:2634485"
 /db_xref="SPTREMBL:O34584"
 /translation="MKFAPMDKVKRFTASHLNKLRLTKKRVPELDDPLLGECWEFEED
 GLKQFWEENYEFVARPKHFNWD"
 complement(3063. .3227)
 /gene="yoqE"
 complement(3063. .3227)
 /gene="yoqE"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13984.1"
 /db_xref="GI:2634486"
 /db_xref="SPTREMBL:O34429"
 /translation="MENVLYGYSTRSRKFCFGEIPEPSKTARKELFNRIIGTDAYKW
 RFEIRKIKRK"
 complement(3281. .4036)
 /gene="yoqP"
 complement(3281. .4036)
 /gene="yoqP"
 /function="unknown"
 /note="similar to phage-related DNA-binding protein
 anti-repressor"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB13985.1"
 /db_xref="GI:2634487"
 /db_xref="SPTREMBL:O34449"
 /translation="MESLYTVIEQNGQLLVDSREVAEMVGKRRHTDRLRSIDGYVAILL
 NAKLSVFEFLESTYKDATGRSLKHFLTRKGDVANKMTGAKVLTFTAQYYSKFE
 MERALKARPSLIDTYLDNEDERAIATVETKAKRELOQLTAEAPVEYDRELNTD
 GLMKIGQAKAIGIKMGQNNLFRLENKVLIDGTNKNAPYOKYVERGFFQVKTQET
 SVGKITITLTPKGADEFIVDLLKKHGHKREIAS"
 complement(4077. .4484)
 /gene="yoqC"
 complement(4077. .4484)
 /gene="yoqC"
 /function="unknown"
 /codon_start=1

```

/transl_table=11
/protein_id="CABL3986.1"
/db_xref="GI:2634488"
/db_xref="SP:PREMBL:O34834"
/translation="MEVNYKTRQREKIRNGIPYDELDTOMHILIDLINFKIGLKTRH
CFCGKPYEIQWFEVNIKEDQILELAELAGREWKGLQISFKWARFSPIMFNWS
LVLSKRPNEDPNKYRLRSVEEFESYAANK"
complement(4491..4829)
/gene="yocqB"
CDS
complement(4491..4829)
/gene="yocqB"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CABL3987.1"
/db_xref="GI:2634489"

```

Query Match 5.4%; Score 64.2; DB 1; Length 213680;

Best Local Similarity 54.4%; Pred. No. 2.1e-08;

Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

QY 335 GCCGGTCATGCACACATCCCCACGGGTACAGCGTCCCTGTACATTCCTCTCAAT 394
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 160924 GCCTGTTGCGCGCTAAATCACGACATCCACCAAAATCCCGCTACATGTTTAT 160983
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 395 GACATCAGGATCCGCGCGTCTCACTGGGGATAACGGCAGCGGAGACTGACGCTTC 454
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 160984 AACCTCAGGATACCGCAATGTTGTTCCATACAGGCACCTCGGACCGCTTC 161043
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 455 AGCCAGTACCATACAAACGCTTCATTTCCGAGGAGCATGACACACATCGCATCCG 514
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 161044 AAGCAGGACAAAGGCAAGCTTCTTTTTCAGATPAGCAGCTTCAATCGTAAAGA 161103
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 515 GTAGACCGGTACGCTGGAAAGGCGACCTGCCATTAACACATCTCCGCTCATTC 571
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 161104 ATAAAGATCTTCAACACGCTTGTGATTTCCAAGCATTAAGACTTGGTCTCCAAGCC 161160
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 20

```

AE003927/c 13334 bp DNA linear BCT 15-JUN-2001
LOCUS
DEFINITION
Xylella fastidiosa 9a5c, section 73 of 229 of the complete genome.
ACCESSION
AE003927 AE003849
VERSION
AE003927.1 GI:9105783
KEYWORDS
Xylella fastidiosa 9a5c.
SOURCE
Xylella fastidiosa 9a5c
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xylella.

```

```

REFERENCE
1 (bases 1 to 13334)
Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
Alvarenga,R., Alves,L.M., Araya,J.E., Baia,G.S., Baptista,C.S.,
Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.,
Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carrier,H.,
Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M.,
Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H.,
Facincani,A.P., Ferreira,A.J., Ferreira,V.C., Ferro,J.A.,
Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R.,
Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A.,
Ho,P.L., Hohnsels,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P.
and Marino,C.L.

```

The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis
Nature 406 (6792), 151-157 (2000)

JOURNAL

MEDLINE

PUBMED

20365717

10910347

2 (bases 1 to 13334)

Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,

Alvarenga,R., Alves,L.M.C., Araya,J.E., Baia,G.S., Baptista,C.S.,

Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S.,

Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carriaro,D.M.,

Carrier,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R.,

Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H., Facincani,A.P., Ferreira,A.J.S., Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S., Gomes,S.L., Gruber,A., Ho,P.L., Hohnsels,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuranae,E.E., Laigret,F., Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A., Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N., Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L., Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C., Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A., Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G., Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C., Palmieri,D.A., Paris,A., Pelxoto,B.R., Pereira,G.A.G., Pereira Jr.,H.A., Pesquero,J.B., Roberto,P.G., Rodrigues,V., de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V., Sawaasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M., Tsuchioka,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.

Direct Submission
Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP 13083-970, Brazil

```

FEATURES
source
1..13334
/organism="Xylella fastidiosa 9a5c"
/db_xref="taxon:160492"
/clone="9a5c"
96..541
/gene="XF0872"
96..541
/gene="XF0872"
/notes="similar to SP|P21364 (percent identity: 30 %/query alignment coverage: 87.3 %/subject alignment coverage: 74.5 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/protein_id="AAF83682.1"
/db_xref="GI:9105784"
/translation="MVSCRRHRPRATKWEAQAQNTVKEGDAVPTLSVTYYINDNVG
FELWITKLLSYTAKTDASGKIGNVTKPTLALSAQYHFGQANVFEPFGVYGSQAK
SKLTALGSTDRVNLGNFRGAIGTIGLDMVDSNWFARLDARTITRFSNSKALSSSL
GLDQNRLLDPWSVFGIGARF"
complement(747..1532)
/gene="XF0873"
complement(747..1532)
/gene="XF0873"
/notes="similar to GI|4156104 (percent identity: 43 %/query alignment coverage: 102.7 %/subject alignment coverage: 98.9 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/protein_id="AAF83683.1"
/db_xref="GI:9105785"
/translation="MKHFLICVATLLFVCGSPDSNHDRLSTAATVPAHETLEFVVP
TLAKOGTOLDIRVFNVDQVQAOQIDVNFQTEPYLOAYNRSRNTHLITVVGVR
IEFGVASYRTALASIPKGAETAIPDPSNNSRALILHOGACILKLPKMLATQR
DILNPLKFRFREDLAMLPRVLDQIDVLINTNYALSAGLKPNQDALIENKDKSPY
NYLVSREDNKDFRVRKALAKALFSPVKDFIDKKYGSAILPAF"
complement(1626..2333)
/gene="XF0874"
complement(1626..2333)
/gene="XF0874"
/notes="similar to GI|6459112 (percent identity: 48 %/query alignment coverage: 91.9 %/subject alignment coverage: 99.1 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"

```

gene
CDS

gene
CDS

Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. .14243
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
complement(102. .767)
/gene="MT2233"
complement(102. .767)
/gene="MT2233"
/note="similar to GB:AL123456; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="IS1558", transposase"
/protein_id="AAK46518.1"
/db_xref="GI:13881921"
/translation="MRSKIPDQALREDDHDMCRHLAHLQDLDMICALDEQ IEQMHFPCARRELIAIPGIGVGSATVISEIGADPAWFPFAEHLASWRCPCGNH ESARKHGGARTGNHQLPVLVECAVAATDGYLREYRQVRKFGFRSPAANK AIIAFAHKLIVIIHWLATGRPYQDLGADYFTTRMDPKERRRLVAKLEAQGLGVTL PAA"
complement(1152. .2540)
/gene="MT2234"
complement(1152. .2540)
/gene="MT2234"
/note="similar to GP:2687633; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phospho-2-dehydro-3-deoxyheptonate aldolase, putative"
/protein_id="AAK46519.1"
/db_xref="GI:13881922"
/translation="MNVVDIPIDQLPSLPPLPTDLRLDAALAKPAAQQTWPADR ALAMRTLVESPVPVTSVLRVQLAQVAKGAPLLQGGCAETFMNDTEPHIGN VRLALQNAVLTICASMPVVKVARIAGYAKPRSDADLGLSYBGDMINGAPDAA AREHDPRLVAYANASAMNLVRALTSGLSLHLVHDNREFVTSAGARKEALA TEIDGRFRSACGAVADNLTQTAETIASHEALVLDYERAMRLSRDGEDGPPQLDLSA HTVMIGERTRIDGAHFAFAQVIANPVGVKPMPELAVYERLDPHKNKCRLTLL VSRMGNKVRDLLEPIVEKQATGHQVWQCDPMHGNTHESSTGFKTRHFDRIVDEVQ GFEEVRAHLPKGGIHVEITGENVTCLGGAQDISETDLAGRYETACDPLNTQOSL ELAFLVAEMRED"
complement(2631. .3137)
/gene="MT2234.1"
complement(2631. .3137)
/gene="MT2234.1"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK46520.1"
/db_xref="GI:13881923"
/translation="WRVYDTEFDTEGHTTIELISIGVVAEDGREYVAVSTEFDPERAG SKWRTVLKLPKPPASOLWRSRQOIRLDLEFLRIDGTDSTELWAWGVAYDHALCOL WGPALPPTVPTRELRQLWEDRGCPRPDRDVEDALVDARDQLRRLITSTD DAGRGAAR"
complement(3147. .4085)
/gene="MT2235"
complement(3147. .4085)
/gene="MT2235"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK46521.1"
/db_xref="GI:13881924"
/translation="WVALYGACISQGGRSFLEVFWHLQHDIVDRGLPLCLCLVAFV

ITFLVTRSFVRFHRRRAADGRPARWQPNRHVIGSVHIIHHVAGVVLVMSGLFLVTL
SYDGEPEFTTAAISIFGGAALVLDDEYALILHLSDVWEEDGRTSDVAFVAVAGL
LIMGLHPLIFFLPROGAAVVLQTLTLAGLVLTPLAVVVLKGVMTGLIGMFVVV
LLVGCASRLSRPHAPWARYTRHPEKMRALQORERTWRPVRRIKLWLQVVIAGTPR
HPDERAVDAQLDQVPPAPPERTAPILISGSVWSD"
4122. .5405
/gene="MT2236"
4122. .5405
/gene="MT2236"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK46522.1"
/db_xref="GI:13881925"
/translation="MSANRAPEVSGRLGRRVLMCLLWLAGVALGYVAVRLFHTPYR
TIDIVQMGAARWLDGRPLYGGVLFHTPIGILNLPFTYPPLAAVLFPFAMLOMPAAS
VAITVTLVLLTASTAIVLTGLDAMPTSRYPAPARLRLAVLIVAPATWLEPIS
SNFAGQINVLMTLVIVDCPPRTWPGMLGIGIAKLTAPVFLYLLLRDGRA
ALTALASAVATLLGLFVLAWSWEYTHHTDRIAGAAINTDONAGLARLTIG
DDERFALWVAGSLVLAATIWAMRRVLRAGETFLAVICVAFGLVVSFVSWSHVWM
LPALVIGLWRRNRNALMSLAGVLMRWTPIDLLPQHRETTAVVWRQLAGMSYV
WVALAVIVAGLTVTARTPQBSLTRLGLTPAPTAS"
complement(5406. .6149)
/gene="MT2237"
complement(5406. .6149)
/gene="MT2237"
/note="identified by match to PFAM protein family HMM
PF01553"
/codon_start=1
/transl_table=11
/product="acyltransferase family protein"
/protein_id="AAK46523.1"
/db_xref="GI:13881926"
/translation="MWYLFKVIEMGPLFLLGRPKVEGLEYPSSGPAILLASNHVAV
ADSEYLPVVRRIWFLKSEYTGTLGWLNRWFYSVSGOVPIPIVAMIGTNVNNPP
TAVLLGQKLLGMPGEGTSPDGRLYKGTGLRLAHLTGVPVPIVAMIGTNVNNPP
SKVDGRFAGVTFVFGKPMDFSRPEGLAGNHFIERAVTDEVIIELMGLSQEYVDIYAA
SVKDGAGNAGAGANPNSTDAARIPETAAG"
complement(6235. .6639)
/gene="MT2238"
complement(6235. .6639)
/gene="MT2238"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK46524.1"
/db_xref="GI:13881927"
/translation="MADVSGAHTDVRPELRKLAQAATLDGIDPAVRVAAVAAVSGGPGT
GKQQVWCPLCALAALVTGEQHPHLLVIADHSLALLEVIRAIIVDDIDRSKPPPEGPP
GGGTGAGGENTNGESMKSHYQAIPTVEE"
complement(6627. .7769)
/gene="MT2239"
complement(6627. .7769)
/gene="MT2239"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK46525.1"
/db_xref="GI:13881928"
/translation="MLVSTDQAHSLGDVLGIATVPVPGQDPPRVVLAIDPEAGGGFLD
ALADTLALLEGRLHVETLDRFPFSGSELSSIAPEELCALPGIQVIGLHVGELAA
ARRWRIVVDCASTADRLMTLPATGLYVERAWRHRLRSIGADGSRSAVLAELE
RIRASVERLSTLLDGAHSAHLVLTPEVVAARAVTRTGLSALMGVRSVEELVNLQ
VODENYSLPDHPAFHYAERIGQRAVLDOLDATIGDVALVIVPHLAGEPIGPKA
LGGILDSARRRQGSAPPGLQPIVDLESGLSIYRLRLALPQLDPLGTILGRADD
LIIVSAGMRRRRLASVLRCTVDAHLRGELTVFRFPNPEVWPT"
complement(7886. .8320)
/gene="MT2240"
complement(7886. .8320)
/gene="MT2240"
/note="identified by match to PFAM protein family HMM
PF01553"

STAFF

Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.

CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or tgg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

```

1: 123456
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
<1...988

```

/strain="H37KV"
/db_xref="taxon:83332"
/clone="Y270"
22587..>23400
/organism="Mycobacterium tuberculosis H37Rv"

```
unit
/complement/115 1469)
/clone="Y190"
/wm_xref="taxon:63332"
```

```
complement(115..142)
/note="Inverted repeat at end of IS6110;
CAGCTCCGGCACTCACCGCGCGCGCTTCA"
```

```
/gene=RV2167C  
complement(157..>1197)  
/gene="RV2167C"
```

```
transposase: RASIA best: IRAG9_MIC10 P19774 putative
transposase for insertion sequence (identical)"
/codon start=1
```

```

/product- hypotrichal protein RV218/C-
/protein_id="CAA17494.1"
/db xref="GI:3242294"

```

GHREGDGLRWGVESICTQLTELGVPIAPSTYYDHNINRPSRRELRGDELKEHII
AANYGVVGARKVM.TINREGIEVARCTVERIMTKIGISCHWIKKARRETAIDQDZ
/TRANSITION--RECTGGMANTANGENDEADRVAFURGATFASILULIRF

NALAEITNGLYKTELKPKGKPRSIEDVELATRWVDWFNHRRLYQYCGDVPVPV
 AVAARORPAAAG*

/note= "fragment designated v021: does not represent a

gene complement(8425. .8814)
/gene="MT241"

```

/genre= Mizzai
/note="identified by Glimmer2; putative"
/codon start=

```

```

/product= conserved hypothetical protein
/protein_id="AAK46527.1"
/db xref="GI:13881930"

```

gene
8971. .10782
RVAGKKMAFEVRPTVLSRPIGVSPVT[#]
RVAGKMAFEVRPTVLSRPIGVSPVT[#]

```
/gene="MT2242"  
/note="similar to Gp:4539200; identified by sequence
```

```

/second_start =
/transl_table=11

```

Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

10894 TTCTGGGCGGCGGCCCATCGGACGGCCGATCCCGATCGATCAGCAACTCGGGCCAC 10955

10954 GGGTCGGCCACCCGGTCCACCGACCTACCGTCGACCACTAGCCAGTCTTGTGTGCTG 1101

11014 CACCGTTCCGGGGCTCCGGCAGAAATGCCGGCATTACCGGCACGCCGGCGGGAGGC 1107

11074 TTGGAGGACACGATGCCAAGCCCTGACGTCC 11107

MTV021 2340 bp DNA linear BCT 03-AUG-2001

AL021957.1 GI:3242293

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;

AUTHORS
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eigelmeier, K., Gas, S., Barry III, C.E.,

Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A.

1. **Introduction**


```

physical clone"
complement(1092..1418)
/genes="Rv2168c"
CDS
complement(1092..1418)
/genes="Rv2168c"
/notes="Rv2168c", len: 108. IS6110
transposase. FASTA scores: sptri008155|008155 HYPOTHETICAL
12.0 KD PROTEIN (108 aa) opt: 697 z-score: 903.5 E(): 0;
100.0% identity in 108 aa overlap. Tbpase score is 0.928"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2168c"
/protein_id="CAAL17472.1"
/db_xref="GI:2911088"
/db_xref="SWISS-PROT:Q50686"
/translation="MSGSSRRVPPELRERAVRMAEIRGQDSWAALSEVARLLGV
GCAETVRKVRQAVDAGAPGTTTIESAELRLRDNALRRANILKTASFAAE
LDRPAR"
complement(1141..1146)
/genes="Rv2168c"
/notes="DRAI:TTTAAA. Site linking fragments L-T"
complement(1442..1469)
/notes="Inverted repeat at 5'-end of IS6110"
complement(1563..1967)
/genes="Rv2169c"
complement(1563..1967)
/genes="Rv2169c"
/notes="Rv2169c", len: 134. Unknown. Tbpase
score is 0.934"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2169c"
/protein_id="CAAL17473.1"
/db_xref="GI:2911089"
/db_xref="SPTREMBL:Q53503"
/translation="MPLSDHEQMLDQIESALYAEDEPKFASVVRGGFRAPTRARLQ
GAALFIGLMVSVAFKFTMIGSPILSVFGVFMGGVYAITGPRLSGRMDRGG
SAAGSRQRTKGAGGSFTSRMEDFRFRFDE"
complement(1973..1979)
/notes="possible RBS, aaggagg, for Rv2169c"
2217..2223
/notes="possible RBS, aggagg, for Rv2170"
2233..2853
/genes="Rv2170"
/genes="Rv2170"
/notes="Rv2170", len: 206. Unknown. Tbpase
score is 0.905"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2170"
/protein_id="CAAL17474.1"
/db_xref="GI:2911090"
/db_xref="SPTREMBL:Q53504"
/translation="MAIFLIDLPDMERRLGDALIVYVDMMRYPRGTETLRAPWLE
HIRRQQAQAAVFAEQAQEAADTALPSAELSNAPMLGVAYGYPAGQWQOQ
VVLQSGRPRLAIRMYSYFELTFLHLPRAQGLGEALRLRLAGDEDNVLL
STPTNGEDNRWRLYRLGFTDIIRGYHFAGDPRAFAILGRTLPL"
2938..2942
/notes="possible RBS, aggagg, for Rv2171"
2949..3632
/genes="lppM"
2949..3632
/genes="lppM"
/notes="Rv2171", len: 227. Probable
lipoprotein contains putative signal peptide and
appropriately positioned PS00013 Prokaryotic membrane
lipoprotein lipid attachment site. Has hydrophobic stretch
at C-terminus and also contains PS00225 Crystallins beta
and gamma 'Greek key' motif signature. Unknown but similar
to . Tbpase score is 0.895"
/codon_start=1
/transl_table=11

```

```

/product="lppM"
/protein_id="CAAL17475.1"
/db_xref="GI:2911091"
/db_xref="SPTREMBL:Q53505"
/translation="MARTRRGLAIAMLMLVPLATCCLRVRSITISPPDLVSGEI
IAAKPRNSKDTGFDALGDVDFPSQKVAVSNDYGVGSQVSDLFALFELPOLANMN
DAAGVNLURLNRNGNIVILEGRADLTVSVDADVELTVAPPAVAVTNGDRIEPEVV
QMKLPQGVVSTMSAQARYTDPNTRSFAGTGWLGIAAFAAGVAVLANWIDRDRSPRL
TASGDPPTS"
2991..3023
/genes="lppM"
/notes="PS00013 Prokaryotic membrane lipoprotein lipid
attachment site"
3162..3209
/genes="lppM"
/notes="PS00225 Crystallins beta and gamma 'Greek key'
motif signature"
complement(3629..4534)
/genes="Rv2172c"
complement(3629..4534)
/genes="Rv2172c"
/notes="Rv2172c", len: 301. Unknown. Tbpase
score is 0.905"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2172c"
/protein_id="CAAL17476.1"
/db_xref="GI:2911092"
/db_xref="SPTREMBL:Q53506"
/translation="MTLNTIALELPPNLEGGKERAIEDARKVQVYSNASGLGRIRH
VMPGMIADDDDRIPMPQPKLDVDFWSIIKPELAGVHGLCTQVAFMDEPSLHRRLV
DLSDAGMEGIYFVGPRTMDGEGSVAPTALSLYQLVANRGVIVPTPDGQGRLL
NFKCSRGATYGTQLLYSDAIVGFLRFPARTTEHREPLLISFGVPKVEIRGLINWL
IODPCNAVADEQAFVOKLAGSEPARRRRLMVDLYKRVLDGVADLGPPLSLHLEATYG
VSAARFETFAEMLAYWSAEPKPD"
Query Match 4.8%; Score 57.2; DB 1; Length 23400;
Best Local Similarity 54.2%; Pred. No. 2.5e-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTCTGGCGTGGTGCATCATCCGGAAGCGTTCCGGTGGGATAAAAATCCGCAG 331
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20072 TTCTCGCGCGCGCGCCCATCCGACGCGCGCGATCCGATCGATCAGCACTCGGCCAC 20131
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 TCCGCCGGTCCATGCAGACATCCCCACCGGTAAACAGCGTCCCTGTTCACATTTCTGTG 391
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20132 GCGTCGCGCACCCCGTCCACCGACTACCGTCGACCACCTAGCCCGAGTCTTGTGTGCTG 20191
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 AATGACATCAGGATCCGCCCGTCTCACTGCGGATAACGGCACCGCGAGACTGACGC 451
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20192 CACCGTTTCCGCGGTCCGCCAGAAATGCCGCGGATTACCGGCACGCGCGCGGAGGC 20251
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 452 TTCAGCGAGTACCATACCAACGCTTCATTTTCC 485
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20252 TTCAGGACACGATGCCCAAGCCCTCGACGTCC 20285
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 23
AF498417/c
LOCUS
DEFINITION
Pseudomonas aeruginosa serotype O6 putative O-antigen biosynthesis
gene cluster, partial sequence.
ACCESSION
AF498417 AC104736
VERSION
AF498417.1 GI:20560066
KEYWORDS
SOURCE
Pseudomonas aeruginosa.
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
REFERENCE
1 (bases 1 to 14734)
Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutyavin,T.V.,
Ivey,R.G., Zhou,Y., Kaul,R., Clendenning,J.B. and Olson,M.V.
AUTHORS
Genetic variation at the O-antigen biosynthetic locus in
TITLE

```


Pseudomonas aeruginosa
J. Bacteriol. 184 (13), 3614-3622 (2002)
22053227
12057956
2 (bases 1 to 14734)
Raymond, C.K., Sims, E.H., Kas, A., Spencer, D.H., Kutyavin, T.V.,
Ivey, R.G., Zhou, Y., Kaul, R., Clendinning, J.B. and Olson, M.V.
Direct Submission
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
On May 13, 2002 this sequence version replaced gi:17975265.
Location/Qualifiers
1..14734
/organism="Pseudomonas aeruginosa"
/serotype="08"
/db_xref="taxon:287"
<1..104
/note="ORF 1; C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027812.1"
/db_xref="GI:20560067"
/translation="EKDAMELRKQEVESAGPTTIGDLIRAQMNQC"
241..525
/note="ORF 2; hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027813.1"
/db_xref="GI:20560068"
/translation="MKSELIERIVHQGLSAKDVELAIKTMLEQMSQALWGDRIE
IRGFGSFLSHYRAPRVRGNPKTGESVRLDQKFPHPKPGKELDRVNEP"
551..805
/note="ORF 3; hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027814.1"
/db_xref="GI:20560069"
/translation="MQWLGRILFAVFLVVALVFLVLENSVSLAFLGWHPTDGP
LSLFLAFLTGLGGLGALLRGRQRKSLRSESTL"
1190..2254
/note="ORF 4; wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027815.1"
/db_xref="GI:20560070"
/translation="MTEKNPSPYKAKSGEDFLIEIMSGLSLRLIALLAVLPVILAI
VILFTPKVTEARVSTLPPLSLDIAGNLSRSEGLPFAFANGVSIYFVRLQSEETR
RQFRDVIYPSLEEGASVGRDQLYALLMNSVTIQKPSKEDSERYTVVIERHDPQAA
ENGEKTIAGVRKSMQEMQEVRRFOVKRNRIQQQIDILRETKAKARDRIISRLKEA
LRVAELRNPLIEGOVDQQLSSIMEGLMYMGAKAIKRAEIKTLQDRVSDPPTIP
ALRGLOEKYALLSGLTILNDKVAFRQDGTIEVDDPVRPRKGIIVLILSLMAGGFLV
AVAILALNVRGFRFRVDA"
2364..3632
/note="ORF 5; similar to UDP-glucose/GDP-mannose
dehydrogenase"
/codon_start=1
/transl_table=11
/product="wbpo"
/protein_id="AA027816.1"
/db_xref="GI:20560071"
/translation="MKDLKVVVGLGVGLPLAVEFGKKRTVGVGFINOGRIARELROG
IDSTLVDAAELAESELSFTNLQLOKCNVEIVTPPTIDSEHKOPDLTPLYKASES
IGVLKGDIVIESVTPGATEDECPVLEKFSGLNFEDFAGYSPERINPGDKHEH
INELAIIFNRMGDTEAVLKAAGTKWNFMFPRGLGVGHGICGVDPYLLTHKAQSIGVH
PEIILARLMDGAGYVWSQVKAMLRRIHVDGARVLLMGLTFKENCPLDKNTKV
DIVRELAENIOVDVDPVWSAEDAMHEYGITVPGPSHGAYDGIILVAHSEKNGK
AENIRLGRKAEHVLYDLKYLDEEDKSDLRL"
3670..4695

/note="ORF 6; similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/product="wbpp"
/protein_id="AA027817.1"
/db_xref="GI:20560072"
/translation="MMSRYBELKELPAQKVMITCVAGFIGSNLLETLLKLDQKVV
GLNFRATGHQRNLDSEVLSERKOWSNFKFIQDIRNLDDCNACAGVYVHLQAALG
SVFNSINDPITSNATINDGFLNMLIARDAKQVSTYAASSSTYGDHPGPKYVEDTIG
KPLSPYAVTKVVELYADVFCYGTIGLRYFNVFGRQDPNGAFAAIVPKWTSSM
IOGDVYVINGDGETSRDFCYIENTVQANLLAATAGLDARNOVNIACVGPSTSLNOLFF
ALRDGLAENGVSYSREPEVYDFREGDVRHSLADISKAALLGATAPKYDVSAGVALAMP
WYIMFLK"
4692..4889
/note="ORF 7"
/codon_start=1
/transl_table=11
/protein_id="AA027818.1"
/db_xref="GI:20560073"
/translation="MIFTVLDESSVVSMSGAIFAFESRLVLQVLNIRGACDSLFS
RSTFSMIVFLFFTHCFLFMV"
4896..6320
/note="ORF 8; similar to Polysaccharide biosynthesis
protein; potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AA027819.1"
/db_xref="GI:20560074"
/translation="MLGKHSIVYFLFKSPAILTVGLSVFTRLLSPGCEGVVSLTII
VVGFLNFFQWALGVGRVLPEDCDQARLLGTARALISFLVSLVIFVTFLLWM
REIGESILYMYGFLCLAQAWHDNLKIQNALQPLTYKMLLIGAGSFFIGLVLV
YFGEVDGLLGLTSLVLAIFQDAWRGWSWALVDKQLEQLTRLFAYGAPLTTLFLFA
FIVNASDRFFIGAFGLGDAVAVSVSYDLAQYSGVTVAHVLAAPLVMEKLSKLSA
POTQDOLRKTFRIFAVVSPAACGLAMVAPISGIMGEFEERGALEIPLISASFL
GALKSFEDYSFOLASATRVQVTVAVSDVVDVFNLLIPERFIGVCAAVSSVMAFSS
AIIISIFLGRVFPMPALPGDKAMKIALSVLLMAVSASFLESFAFLGVKVVVIGGG
VYLAAMIALDVSGMRTFLKSLIR"
6517..7488
/note="ORF 9; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AA027820.1"
/db_xref="GI:20560075"
/translation="MKALLEGLTGIDMHMSALVSAYPARRLDRLPLMLRELLHR
LPQVIGSYSDVILQRELLSTIPLFLTKAPRILDDDAIWLHRRGIAANSIARV
DHLVCGNOYLADYFGGRPTTIPTGVDTLRFLPRERRENVRVIGSGTSGGYRFLY
DIEPLSLFRDHPKWLRLVSDRPPPEKVIPEKVEFIWTEETNEVETAGNDIGIM
PLADDLSRGKCSYKMLLYNACGLPVVSYGMNROVLARGFIGYGVDDGWTESLA
ALVKDPEARVRAGONGHDIITERHYSLDVVDVCDLMAWVIGSVAPGKGR"
7493..9376
/note="ORF 10; similar to Asparagine synthase"
/codon_start=1
/transl_table=11
/protein_id="AA027821.1"
/db_xref="GI:20560076"
/translation="MGIAGFWNITGTLGDNARVARQMAAIIHRRGPDSEGIWYEAP
RAPILVHARLAVLELSPAGSQPMHSCGRVLYNGEYIHLARLARSAGVYTHSWR
GGSDTETLACFAQWGVSTLKLTVGMFALADWROEKTITLARDMGEKPLYGWQON
GVLFASSELKALKEHPLFRGDIIDRDALALFERYVPAPYSIKGICKRAGSVLYLVS
ERSLNTECPAAAYWASNALEALSFPQGTDAEAVDLLESQRTSISDQMSDVPLG
AFLSGGVDSSTVVALMQQSSRPIITFSGIDFEPGYEAVYAKAVAEHIGTDHTELYV
NSKDALDVPISLPKIYCEPDSQIPLIVSGLRQOQTVVALSGDGDDELFGYINPY
NHEPYVWRMLEREPHSMRREASAFODLPLPEKIGLRDVFASRTAEKLYRLNSHR
OFTYPIVGAQHTALLDTPERNWRPVSFQHWAMVDVQGYMPDDILLYKVDRAAMNSL
ETRVPLIDHVRFLAWPLUMKIRNGKWLREVLVYRHVSRELIERPKKGSVPVPS
DWLRGPKLWAESELLDRLOQEGYLDLSRLIRIWNHDLGRDRHRSRLVLMFOAM
LES"
9373..10524
/note="ORF 11; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AA027822.1"


```
IGKVLKKDIVIESTVYPGATEEDCVPLEKFSGLRNFNEDEFAGYSPERINPGDKEH
RVSSINKVTSSTPEAEALVDSLIYREITITAGTHKASSIKVAEAAKVIENTQDLNIAL
INELAIIFRMGIDTAEVLKAAAGTWNFMPTFRGLVGGHCIGDVPYILTHKQSIGYH
PETILLAGRLNDMGAYVVSQVLKAMLRRIHVDGAVLLMGLITFKENCPDLRNTKVY
DIVRELAENYIOVDVDPWVSADAMHEYGITPVGTPSHGAYDGIILAVAHSEFKNMG
AMNIRKLGAEHVLVDLYKLLDEKSDLR"
4370..5395
/gene="wbpP"
CDS
4370..5395
/gene="wbpP"
/codon_start=1
/transl_table=11
/product="WbpP"
/protein_id="AAF23998.1"
/db_xref="GI:6690131"
/translation="MMSRYEELRKELPAQPKVWLITGVAGSIGSNILETLKLDQKVV
GLDNFATGQRLNDEVRSLASEKQSNFKFIQODIRNLDCCNACAGVDYVLHQAAIG
SVPRSINDPITSNATINDGFLNMLIAARDAKQSFYTAASSTYGDHGLPKVPTIG
KPLSPYAVKYVELYADVFSRCYGFSTIGLRYFNVEGRQDPNGAYAAVPIKWTSM
IQGDDVYINGDETSDFCYIENTVOANLATAAGLDARQYVNIAGVGTSLNQJLF
ALRDGLAENGVSVTHREPVYVDFREGDVRHSLADIKAALKLGYAPKYDVSAGVALAMP
WIMFLK"
5392..5589
/gene="wbpQ"
CDS
5392..5589
/gene="wbpQ"
/codon_start=1
/transl_table=11
/product="WbpQ"
/protein_id="AAF23999.1"
/db_xref="GI:6690132"
/translation="MIFTLVRLDESSVMSGSAIFAFESRLVLQVLNIRGACDSLFS
RSTFSMIVFLFTTHCFLEFWM"
5596..7020
/gene="wzx"
CDS
5596..7020
/gene="wzx"
/codon_start=1
/transl_table=11
/product="Wzx"
/protein_id="AAF24000.1"
/db_xref="GI:6690133"
/translation="MLGKSLYFLFKSPAILTLVLGVSFTRLLSPGEGYVSLTII
VGLFNTVFLQWALGVGRYLPECSDQARALRGTAARLISLVLSVLIIFVTFLLMEW
REIGFSIILYMGVFCIAQAWHDLNKTQNALIQLPTLYGKMLIKAGSFFGLVLLV
YFGVGDGLIGLTLVSLVATIFFQDAWRGVSWALVDKQLTRLFAYGAPLITFLFA
FIVNADREFFIGAFLGDAAGVYVSVDLAQYSGVTAVSVVHLAAPLVNKLKSGGL
POTQDLRFTTFFIYAVVSPAAGLAMVAPELISGSTMGEFREGALKIIPILISLAF
GALKSFYDYSQLASATRVQVTVAVSAVDVFNLIILIPEGIVGAAVSSVMAFSS
ALISIFLGRVFPALPGKDMKIALSVLLMAVSVASFSLAESAFFGLVVRVVLGGG
VYLAAMIALDVSGMRTFLKSLIR"
7142..8188
/gene="wbpR"
CDS
7142..8188
/gene="wbpR"
/codon_start=1
/transl_table=11
/product="WbpR"
/protein_id="AAF24001.1"
/db_xref="GI:6690134"
/translation="MTAELPAVVRVRAFTQGRSVPSRFRVKALLEGIGTRGIDMHSE
ALVSAVPPARLDRLPWLIRLRLPQVIGSYGSDVILQIRELLSTIPTLEFLTKAP
RILDVDDATLWRGLAANSIARRVDHIVCGNOYLADYFGQFRPTTIPTGVDITLRF
LPRRERRENRIGWSTSGYRFLYDIEIPLSHLFRDHPKMLRIVSDRPPKVP1PA
EKVEFTWTFETNEVETIAGMDIGIMPLADDLRSRCKSYKMLLIYMACGLPVVYSEGM
NRDVLARGIGVAVDDEGRYESLAALVNDPEARVRAGNGRDIIRRHYSLDVVCDLW
AMVIGVAPGKTR"
8193..10076
/gene="wbpS"
CDS
8193..10076
/gene="wbpS"
/codon_start=1
/transl_table=11
```

```
/product="WbpS"
/protein_id="AAF24002.1"
/db_xref="GI:6690135"
/translation="MCGIAGFWNITGTLGDNARVAROMAAAIHHRGPDSEGIWEAP
RAPILVHARLAVLELSPAGSQPMHSDGVRVLYNGEINYHLALRLARLSEAGVTHSHR
GSDTELLACFAQMGVESTLKLTVGMFALALWDRQEKITILARDMGKPLTWGQNR
GVLPFASSELKAKEHPLFRGDDIDRALALFLRYGYPIFYITKGLKURAGSILVLS
ERSLNETCEPAAYSANAIEEALSNPFGQDAAVLLSOLRTSISQMSVDVPLG
AFISGGVDSSTVALMQQSSRPITFISGIDEPGDEAVYAKAAVAEHGICDHTELYV
NSKDALDVIPSPLKIYCEPFGDSSQIPTLISGLARQOVTVALSGDGLFGGYNPY
OFTPVWRMLERPHSMRFPASAFODLPLPEKLGKLRDVFASRTAEELFYRLNSHWR
NHEYVPIGAGHTALLDTPRMPRVDSFQHMWMDVQGMEDDILVKVDRAMAHSNL
ETRYPLIDHRVFEELAWRMLHMKIRNGKRWLLREVLIRHVSRELLERPKFSGFVPS
DWLGRPKWEASLDERRLQDEGYLDSRLIRKRWHLHAGRDHRSRLWSVLMFOAM
LES"
10073..11224
/gene="wbpT"
CDS
10073..11224
/gene="wbpT"
/codon_start=1
/transl_table=11
/product="WbpT"
/protein_id="AAF23993.1"
/db_xref="GI:6690126"
/translation="MTTVKVIHVIAGLKKVGGAEMLMKRLIETQNGSGSEPEHSIISL
DLGEFGGLIEAGISVDVIGMTSMRDMRPVLLRLIWIIFRRRPDIYOTVMYHSDLLGG
LAARMAGIGIIGWVRTTDLQEGKSTTVLRKVCWLSGFLPKYICVCAASRRSHI
AVGYNACRLMVLIPNGPDLRLQATDQSGAISRGAGIEASDIVIGSLGRPHVPKDHAS
FVAAGALLAPYSLKFLVGLRELLSSNAELORLIRATGYAERFILLGERODVASCLK
AMDFCLHSRTGTFPNVIGEMAMGLPCTITDVGDAAYLLGNDGVVVPALDPNALGKG
IEDLIALDVGRGALGEAAQRISYNSFTMASQSFRMSLYRDLILVRKKA"
11268..12398
/gene="wbpU"
CDS
11268..12398
/gene="wbpU"
/codon_start=1
/transl_table=11
/product="WbpU"
/protein_id="AAF23992.1"
/db_xref="GI:6690125"
/translation="MKPVLLMIVNDPAFMSHRLPVAVGAQAQAGFVHIATRPDVK
KIVSQGFLHHELPLRSRGNKPNFSELYLLVYWRLLWRDPDLVHLTIKVIYGGINA
RLAPVGVAAVSGLGFVFMAGLKACACRACVAVWLYRRALRKLKRLVIFQNPDDRDA
LIGLGAITTEKSVLIRGSDVLTQYQAPSEPTPVYTLAARLLRDKVLEFVEAANI
LQRSYPAHFQLVGLDLPNGNTIEPSELERWRESEGTIECLGRQDIASVFAFSHIVV
LPSYRGLPKVLVEASACRAVYTVDPVGCDAIQADVTGLVLPVRDSALADAIOBL
IESPELRKKMGAAGRALAERDPAIESIVQQHLDIYRALGSGA"
12395..13357
/gene="wbpV"
Query Match 4.8%; Score 57; DB 1; Length 16736;
Best Local Similarity 49.8%; Pred. No. 2.7e-06;
Matches 144; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 234 TGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATATCTTCTTGGCGTGGCGCATC 293
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12355 TGGACGATACTTTTCGATCGCAAAATCACGCTCTGCCAACTCGCCCGCAGCAGCTCCCATC 12296
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 ATCCGGAAGGTTCCCGTGGGATAAAAAATCCGCGAGTCGCGCGGTCCATGTCAGACACA 353
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12295 TTCTTGGCAACTCTTCGAGCTCAATAAGACGCTGAATGTCATCCCGCAGACGACGAGA 12236
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
354 TCCCCCAGCGGTAACAGCGTCCCTGTACATCTTCTTGAATACATCAGGATCCCGCC 413
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12235 TCACGACCGGACAGGAGCGGCCGCGTAACATAGCTGATTCGATCCCGCGAACCGGT 12176
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
414 GTCTCACTGGCGATACGGGACCGCGGAGACTGACGCTTTCAGCGTTACGATACCATCAAC 473
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12175 ACATCGTCTGTCAACAACCGCAGCCGCGCAGCGCTCGACGACGCTTAGGAAGC 12116
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
474 GCTTCATTTTCCGAAGGATGACCACTGGCACTGGCACTGGCACTGGCACTGGCACTGG 522
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12115 CCTTCAGATAAGAGGACGAGACTATATGCGAAGCGGCGGAGACAG 12067
```

RESULT 25
 AP003599/c
 LOCUS 237221 bp DNA linear BCT 28-NOV-2001
 DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 19/19.
 ACCESSION AP003599 BA000019
 VERSION AP003599.1 GI:17134317
 KEYWORDS
 SOURCE Nostoc sp. PCC 7120 DNA.
 ORGANISM Nostoc sp. PCC 7120
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 Kaneko, T., Nakamura, Y., Wolk, C. P., Kuritz, T., Sasamoto, S.,
 Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T.,
 Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A.,
 Nakazaki, N., Shimpou, S., Sugimoto, M., Takazawa, M., Yamada, M.,
 Yasuda, M. and Tabata, S.
 Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120
 DNA Res. 8 (5), 205-213 (2001)
 21595285
 REFERENCE 2 (bases 1 to 237221)
 Kaneko, T.
 Direct Submission
 Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research
 Institute, The First Laboratory for Plant Gene Research; Yana
 1532-3, Kisarazu, Chiba 292-0812, Japan
 (E-mail: kaneko@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/cyanobase/
 Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
 FEATURES
 source location/Qualifiers
 1. .237221
 /organism="Nostoc sp. PCC 7120"
 /db_xref="taxon:103690"
 /note="synonym: Anabaena sp. PCC7120"
 complement(72..734)
 /gene="all5179"
 complement(72..734)
 /gene="all5179"
 /note="ORF_ID: all5179"
 /codon_start=1
 /transl_table=11
 /product="O-methyltransferase"
 /protein_id="BAB76878.1"
 /db_xref="GI:17134318"
 /translation="MTRTLGIPTPNLYLLSVLSREPEILAQRLQETALQPMGRMOI
 APQGGFMALLVLLGAKKTLGVGFTGYSSLIVALPAEGLKLVACDINEEFTALQ
 RYQKAGVDHKIDLYLAPALETLDKLLVAGEAETDFAFIDADKSNYDNYERSQLI
 RSGGVTAIDNVLWSGVADPEIQDNRTQIRAFNHKLQDQRTILSLPIGDLTLVR
 KN"
 867..1568
 /gene="alr5180"
 867..1568
 /gene="alr5180"
 /note="ORF_ID: alr5180"
 unknown protein
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76879.1"
 /db_xref="GI:17134319"
 /translation="MTLIGATDPRDIPSYISDAARYLRIPAGTIRSWTGRHYPIS
 NGNFRPLIQCNAPRLLSFTNVEVHLRAIRKHQIDILGKVRDALDFIDEQFOI
 SHPLARETLGDVLFIERYGLSINASKSVKTELDAFNALRIEPTDGTGLAIKLY
 PFTFRSHEDNPREVVYDPRIAFGRLVIGTGISTRVLAERYQAGESIDELAYDYDCDR
 LMLEEARCELPATA"
 1697..1996
 /gene="alr5181"
 1697..1996
 /gene="alr5181"
 /note="ORF_ID: alr5181"
 unknown protein
 /codon_start=1

/transl_table=11
 /protein_id="BAB76880.1"
 /db_xref="GI:17134320"
 /translation="MEWPEVGKRGVVLTKDGKISNNRLERIAVARAQIKWFTFASQ
 SLSGEMAGILLQAIYPMKRFVSKHPAFIAKIYRDGHLDMKQAMLEELQEF"
 2453..3310
 /gene="alr5182"
 2453..3310
 /gene="alr5182"
 /note="ORF_ID: alr5182"
 /codon_start=1
 /transl_table=11
 /product="oxidoreductase"
 /protein_id="BAB76881.1"
 /db_xref="GI:17134321"
 /translation="MAEKQTLPQQQQPPCTESKMQPKPOADDAQYRGSGKLKDKYA
 LITGDSGIGRAVAIYAKEGADVAFVYLSEHDDAEETKNLVEQGRRAVSGDITD
 EAFQRAIQQTVEFGKLDILINNAEQHPESIEDITKEQLERTFTNIFSMFLTK
 AAIHLKRGSAIINTSVYKYGKPOLLDYSTKGAIIVAFTRSLNSLISKIGRNVA
 APGIWTLPISPTFAEKVETFGQVPMQRAGQPEVAPSYVFLASDDSSYMSQVLLH
 PNGGEVNG"
 3552..3809
 /gene="asr5183"
 3552..3809
 /gene="asr5183"
 /note="ORF_ID: asr5183"
 unknown protein
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76882.1"
 /db_xref="GI:17134322"
 /translation="MSSLKPLQGTVELDCARANAKQGIETAAQYQCYGDDLNRPAFEL
 RKACEMLQVKELNELITDQDMILNLGTGIIAPDTASEL"
 complement(3821..5221)
 /gene="all5184"
 complement(3821..5221)
 /gene="all5184"
 /note="ORF_ID: all5184"
 /codon_start=1
 /transl_table=11
 /product="Mg2+ transport protein"
 /protein_id="BAB76883.1"
 /db_xref="GI:17134323"
 /translation="MTENNLNSTLQDVSRRELRLVTRQMRMLLEAGDLQAKAILV
 VPQADIAEAIEGPEYMHAFRLSKDEAIEVYEDYDVSQERLEELKRSQVRDI
 VQSDSDARLFDLPAKVNNHLEQLSPTEQATALMLGYEADTAGRTMTLEFIAL
 KENMTIGALERIRSLANASMIYVLYTQDARLTGIVSLRELVTSPQETIGEVMT
 RDVLFVNTDQHEVARLIQRYDPLAVPVVDROQLLGIYTVDDVIDILEETPKDIY
 ALGGVQSGGNYFQMDLWEARKRVLWLFVLLITNVGTIISKQEDILTKVVTLLA
 FIPLLTGGNGVGAQSSVTVIRGMNTDEIRSGPLQVIGREAIAGALGMLGSIATI
 WAYLQGRIEVAIAGASLIAISVLASISGSALEFLFRYLRLDLPALMSAPFTTAVDV
 LGVLIYFNLAIRVILRL"
 complement(5361..6890)
 /gene="all5185"
 complement(5361..6890)
 /gene="all5185"
 /note="ORF_ID: all5185"
 /codon_start=1
 /transl_table=11
 /product="mercuric reductase"
 /protein_id="BAB76884.1"
 /db_xref="GI:17134324"
 /translation="MSNSELDRVTRVPDEYNOKLVAIVHPHPWNVNPPQADNYDLVVI
 KAGTGLVVAAGAGLGLKVALIERKLMGDCGLNVGCVPSKTIIRSSRVVGINGW
 NGLVNPISQIDIDFPVAMRMVRVAGISHDNSAERFASLGVDFVLGSGRFASNTV
 EVACKTLKFKKAVIATGATKPAIIGEOAGYLTNETVPSLIQRPKLAIVGGPGIG
 CELAQAFRLGSEVVLHSGSHVNLKEDNDAAQIVQQTLLKEGIRLVLNKAVEVTVV
 TQRLRYFTNGHSDSVTDILVAGRNSVEGLNEAVGVKDYKRRGVNDVYLTQ
 TNPKIYAAGSDICMDKFTFSSVADAAIRIVIKNTLSPFGLGRSLSLVMPWVYTDPE
 VAHVGLYESQDQVETIKIPFSSVADAAIADQEDGFLKIHKKGSDEIVGATIVASHAGE
 MISEVTTAIVNKIGLNLSNVHPYPTQAEAIKKAADYTRTLTLTPRTKLLGLTLTKF
 S"
 7074..7826

gene

/gene="alr5186"
 7074. .7826
 /gene="alr5186"
 /note="ORF_ID:alr5186
 hypothetical protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76885.1"
 /db_xref="GI:17134325"
 /translation="MIPKQETRNKRHLLIGLTAILIVARQENICALLQTLILW
 VQSLGFGPIAYMILYNLATLIFPGSILTLKSGCLFGVFWGVSVVLIAATVCAILAF
 FIGYLSRWYVQIDIKYKPKFMIDSVAKGKIVLLRLSPVFFPNLINTAFGVTIC
 ISKDLTGLGIPGTIMVYIGDSVAGDLAKAGLTQHQPVTPETQITQWIMOGGLGLMA
 TVGVYIITKIAQKALSQKVVTGIIKSQDAE"
 7859. .8011
 /gene="asr5187"
 7859. .8011
 /gene="asr5187"
 /note="ORF_ID:asr5187
 unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76886.1"
 /db_xref="GI:17134326"
 /translation="MLTTDDTKIKHPFLGVNHEEKVEWQETLLQLSLSLEIPKEVASD
 ESHSHS"
 7989. .8657
 /gene="alr5188"
 7989. .8657
 /gene="alr5188"
 /note="ORF_ID:alr5188"
 /codon_start=1
 /transl_table=11
 /product="two-component response regulator"
 /protein_id="BAB76887.1"
 /db_xref="GI:17134327"
 /translation="MNRLLTAEDPRIAIEFKLSQSGTTAVATDAYSATNMALSS
 GFDLIDLGPKGDLVDLEELRGCGENIPVILTARDIQDKVAGFAGADDYLTIK
 PPFIEELLVRKARLQSGSQAMEETIKSGNIVLDLSRKVKIGQDTIELPAREFT
 LAETFFRHQVQLSRQQLLDYRWGVDYDPGSNIVDVYVGYLRKLLGNDLIEIVRGNGY
 RLPT"
 8884. .10497
 /gene="alr5189"
 8884. .10497
 /gene="alr5189"
 /note="ORF_ID:alr5189"
 /codon_start=1
 /transl_table=11
 /product="two-component sensor histidine kinase"
 /protein_id="BAB76888.1"
 /db_xref="GI:17134328"
 /translation="MGQQAQKESQVILFHPDQKSVKNOQNTIAGTFLKHRRGFPWS
 TRPILSWGIIILFFELFPIPIRQALYARVNRVNEEAIEKIEFQELLGTSNIPK
 HQDESIGAIKDKTIDNRKLPPPTKQQLREFFDALGNQLPEDDTFLIALMDKFI
 YKSIPARPKEMDRDAKLIDNAKLTPQKEEYVTPNSQVDSIVYLARVEVIEQGEIM
 VVIATHTAGEERDEVLTLAVIVQSTFVLIFALVLAWLASGKILAPLRLLTQTARKI
 SETDLNORTIDNGBELAELGTTTFNDMDRLQNAFIQSQNFINDACHELRLTPIILRG
 HLELMNDPKIEQETLTIVIDELEMRNFRNDMLLAKAERPDFLLELFDVCTTFEE

Query Match	Best Local Similarity	Score 54.8; DB 1; Length 237221;
Matches 137; Conservative	0; Mismatches 137; Indels 0; Gaps 0;	
QY 275	TCTGGCGCTGCCTGCCATCATCCGGAGCGTTCGGTCGGGATAAAAAATCGCGCACTGC	334
Db 74530	TCGGGGCGCTCTCCCACTGCTCTGCTAATCTTTGGGATGTGAATACACCATTCAGCTT	74477
QY 335	GCCGGTCCATGCAGACACATCCCGGTAACAGCGTCCCGTCCACATCTCTCGAAT	394
Db 74470	TTCTTTGAGTGCCTTACATCTCGATGGGAATAATGAACCACTCTACTCCATCAGTAAT	74411
QY 395	GACATCAGGATCCCGCCGCTCTACTGGCGGATAACGGGACGCGGAGACTGACGCTTC	454

Db 74410 AATATCAGGCGCTGCTGTTATTATATTAATGTTAAATTACAGGAATTTCCACAGAGATGGCTTC 74351

QY 455 AGCCAGTACCATACCAAAAGCGCTTCATTTCCGAAGGCATGACACACACACTGCGCAATCGG 514

Db 74350 TAATAGAACTAGTCCAAACCCCTTCATTTAAAGAAAGAAATACTTAAAAACCGTAGCACTGCC 74291

QY 515 GTAGACCGGTAAACGCTGGGAAAAGGCGACCTGCC 548

Db 74290 GTAATACTGGTTCAGAAGTAATGGGGAAGTGC 74257

RESULT 26
AE008327/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
CDS

AE008327
Agrobacterium tumefaciens str. C58 (Cereon).
of 187 of the complete sequence.
AE008327 AE007870
AE008327.1 GI:15159773
Agrobacterium tumefaciens str. C58 (Cereon).
Agrobacterium tumefaciens str. C58 (Cereon)
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Rhizobium.
1 (Bases 1 to 10197)
Hinkle,G., Slater,S.C. and Goodner,B.
Complete Genome Sequence of Agrobacterium tumefaciens C58
(Rhizobium radiobacter C58), the Causative Agent of Crown Gall
Disease in Plants
Unpublished
2 (Bases 1 to 10197)
Hinkle,G., Slater,S.C. and Goodner,B.
Direct Submission
Submitted (14-AUG-2001) Bioinformatics, Cereon Genomics, 45 Sidney
Street, Cambridge, MA 02139, USA
Approximately 800 bp of telomeric sequence missing from the left
end of the chromosome and 200 bp missing from the right end.
Location/Qualifiers
1. .10197
/organism="Agrobacterium tumefaciens str. C58 (Cereon)"
/strain="C58"
/db_xref="taxon:181661"
219. .1367
/gene="AGR_L_2541"
219. .1367
/gene="AGR_L_2541"
/note="(17900) putative hexose transferase"
/codon_start=1
/transl_table=11
/product="AGR_L_2541p"
/protein_id="AAK89838.1"
/db_xref="GI:15159774"
/translation="WTSALFVSHTEKGAEFLADLVKAGPHSRACPLSGATAED
LAEAGPVMLSAGEKMLSRNNSFGALGAADVMAWQLSREAKHFVDCANSQ
KALFVCAAKLSRRLVILHDIVDTAFATNRRLASLAFARILVAVNSEETGR
OFTVEGAEADKRVIVYNGVDPAAKLHDAGMAARLRAELGLGQPLVGLFGRLSWKQ
AHFLDALAAMEGVQAVIVGGALFGOEAYEARIQEASRLGLDGRVRLGFRSDVPEL
MASMDVAVTSTVAEPFGVVEAMCGRPVAVRGGGVTEIIRGETGLLVPPGDAS
ALAAAGLTSLDPAQLQGSGREDVSDRFSLQETCRSVSALLTEAA"
1396. .2514
/gene="AGR_L_2543"
1396. .2514
/gene="AGR_L_2543"
/note="mannosyltransferase B - Synechocystis sp. (strain
PCC 6803)"
/codon_start=1
/transl_table=11
/product="AGR_L_2543p"
/protein_id="AAK89839.1"
/db_xref="GI:15159775"
/translation="MPAGGRAMRILNMPQSYGKGASGVARTFSLERLLEQFGDD
YILRSPWTRQEPQLSRLEMETPRRIMVFDVLRQAAMLPVLCRRHGIDVLFNY
DFDPSLPGCKRLTLVHDLYFKTIPQIGWRATLTDTDFRLMAGSNRVVCVSEAT
KOLAREYPAATKSLTIHSDSLTSDPQDPDWTASPTAGRYTILAVGNATSNKWFATIG
K"

RESULT 27

AE009285	LOCUS	AE009285	12006 bp	DNA	linear	BCT 20-DEC-2001
DEFINITION	Agrobacterium tumefaciens str. C58 linear chromosome, section 55 of 187 of the complete sequence.					
ACCESSION	AE009285 AE008689					
VERSION	AE009285.1 GI:17741958					
KEYWORDS						
SOURCE	Agrobacterium tumefaciens str. C58 (U. Washington).					
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Rhizobium.					
REFERENCE	1 (bases 1 to 12006)					
AUTHORS	Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F., Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.					
TITLE	The genome of the natural genetic engineer Agrobacterium tumefaciens C58					
JOURNAL	Science 294 (5550), 2317-2323 (2001)					
MEDLINE	21608550					
PUBMED	11743193					
REFERENCE	2 (bases 1 to 12006)					
AUTHORS	Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F., Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.					
TITLE	Direct Submission					
JOURNAL	Submitted (27-SEP-2001) Department of Microbiology, University of Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA 98195-7242, USA					
FEATURES	Location/Qualifiers					
source	1..12006					
	/organism="Agrobacterium tumefaciens str. C58 (U. Washington)"					
	/strain="C58"					
	/db_xref="taxon:180835"					
gene	100..360					
CDS	/gene="Atu3552"					
	100..360					
	/gene="Atu3552"					
	/note="putative; ORF located using Glimmer"					
	/codon_start=1					
	/transl_table=11					
	/product="hypothetical protein"					
	/protein_id="AAL44364.1"					
	/db_xref="GI:17741959"					
	/translation="MTKIVFSTGTGGISDFRGAVKVRCCENLSRVTVNSEFKCRREAIGFIIVDNALIDNIYFISKTWILERLOTETSTF"					
	727..1785					
gene	/gene="Atu3553"					
CDS	727..1785					
	/gene="Atu3553"					
	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"					
	/codon_start=1					
	/transl_table=11					
	/product="conserved hypothetical protein"					
	/protein_id="AAL44365.1"					
	/db_xref="GI:17741960"					
	/translation="WTPKDRNSPVQRLRRIVASNIRLTICIDNLPYVTPDLRL"					

ADPYAWCDGIDIFEGSNENLTKQNPAPRISAKHQAAALYNVHNSPQLSAVPVVAP	gene	ADPYAWCDGIDIFEGSNENLTKQNPAPRISAKHQAAALYNVHNSPQLSAVPVVAP
SVYLTHRALDLKDCADYGNITHPYAGMEHPETPGKLSKSAIAHAAAGPVPVAT	CDS	SVYLTHRALDLKDCADYGNITHPYAGMEHPETPGKLSKSAIAHAAAGPVPVAT
EMGYHTSLATKTLHFHPTGIGIKARYPMRLWSFISGIRRSYIEVLSFAADATNPE		EMGYHTSLATKTLHFHPTGIGIKARYPMRLWSFISGIRRSYIEVLSFAADATNPE
SFGLRLHDLSTPAYEAVRGLLSLCKANROTTSVREVAFLDADAQRLSMPTTRODG		SFGLRLHDLSTPAYEAVRGLLSLCKANROTTSVREVAFLDADAQRLSMPTTRODG
ALLVPVLAWSGQWPALENPPAERVAFAFSVAGHSRVVVAHFRFLDDGVSQQTIAPE		ALLVPVLAWSGQWPALENPPAERVAFAFSVAGHSRVVVAHFRFLDDGVSQQTIAPE
GGQRLSVSDQLTVLEVF		GGQRLSVSDQLTVLEVF
1992..2429		1992..2429
/gene="Atu3554"		/gene="Atu3554"
1992..2429		1992..2429
/gene="Atu3554"		/gene="Atu3554"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"		/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1		/codon_start=1
/transl_table=11		/transl_table=11
/product="conserved hypothetical protein"		/product="conserved hypothetical protein"
/protein_id="AAL44366.1"		/protein_id="AAL44366.1"
/db_xref="GI:17741961"		/db_xref="GI:17741961"
/translation="MRAVAAMIVAHADVNEFTFLPLRALRLGGGVDFLVICGFV		/translation="MRAVAAMIVAHADVNEFTFLPLRALRLGGGVDFLVICGFV
IYASRPYFAAGGRSAFLRLRFLRVPLVFLVRLMAAAVPTIIGAKAFPPPEWSN		IYASRPYFAAGGRSAFLRLRFLRVPLVFLVRLMAAAVPTIIGAKAFPPPEWSN
LVYLVIVISTVAIAHAFVFEVETPANLMLRLDRLITSPRKVAVS"		LVYLVIVISTVAIAHAFVFEVETPANLMLRLDRLITSPRKVAVS"
2611..3537		2611..3537
/gene="Atu3555"		/gene="Atu3555"
2611..3537		2611..3537
/gene="Atu3555"		/gene="Atu3555"
/note="putative; ORF located using Glimmer"		/note="putative; ORF located using Glimmer"
/codon_start=1		/codon_start=1
/transl_table=11		/transl_table=11
/product="hypothetical protein"		/product="hypothetical protein"
/protein_id="AAL44367.1"		/protein_id="AAL44367.1"
/db_xref="GI:17741962"		/db_xref="GI:17741962"
/translation="MPRIHNGSLVGMVASFVSGVSSVANAGTQVETLPPAAPQOQPA		/translation="MPRIHNGSLVGMVASFVSGVSSVANAGTQVETLPPAAPQOQPA
VSAATGAPVTAADSEATFLARVNTNPDGVSISAAELFKQSPQAIYKDLVTIASS		VSAATGAPVTAADSEATFLARVNTNPDGVSISAAELFKQSPQAIYKDLVTIASS
TDAATAVTGKVTSGEPIIYAVSTSAQFSEQLSILSDSGQLPTVVAAQKVGKDF		TDAATAVTGKVTSGEPIIYAVSTSAQFSEQLSILSDSGQLPTVVAAQKVGKDF
SEVVGELALAHATMRNGNTILAAAIQIQTKGPGVPSDLLTAFSNIQCTAAGRAAA		SEVVGELALAHATMRNGNTILAAAIQIQTKGPGVPSDLLTAFSNIQCTAAGRAAA
AAAPAASTPIGGAGSSPASANSIGGSGTAGTEGTAQAGGAPVGGGSGVGGTISGG		AAAPAASTPIGGAGSSPASANSIGGSGTAGTEGTAQAGGAPVGGGSGVGGTISGG
GTSPSTTNFTTINNFFSTGNTTTPVSPV"		GTSPSTTNFTTINNFFSTGNTTTPVSPV"
3642..5918		3642..5918
/gene="Atu3556"		/gene="Atu3556"
3642..5918		3642..5918
/gene="Atu3556"		/gene="Atu3556"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"		/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1		/codon_start=1
/transl_table=11		/transl_table=11
/product="exopolysaccharide polymerization/transport protein"		/product="exopolysaccharide polymerization/transport protein"
/protein_id="AAL44368.1"		/protein_id="AAL44368.1"
/db_xref="GI:17741963"		/db_xref="GI:17741963"
/translation="MLQRHDPMTSDYDKSLTAEFIDIDAALITRQWRVVLAAIA		/translation="MLQRHDPMTSDYDKSLTAEFIDIDAALITRQWRVVLAAIA
VAGIAGLAFAATAVPIYSATATLLIDRNSQIVQOLSTIGGVYDEASLSQVEVLQS		VAGIAGLAFAATAVPIYSATATLLIDRNSQIVQOLSTIGGVYDEASLSQVEVLQS
ETIGLAVDSLKLTENQEFRASTRASLLSIFGTIRSLVNSQWFSPTKKEAVIDDGTL		ETIGLAVDSLKLTENQEFRASTRASLLSIFGTIRSLVNSQWFSPTKKEAVIDDGTL
KRSLDRLLNDLSVKRIGRTYALERTYNSTPVLAQIYNVAVASAYLLDKLSKYEAT		KRSLDRLLNDLSVKRIGRTYALERTYNSTPVLAQIYNVAVASAYLLDKLSKYEAT
RRASDWLSRIAELRQRLDLDLAVOKFRAEHLITTGNGLLSDQOLAESNALILA		RRASDWLSRIAELRQRLDLDLAVOKFRAEHLITTGNGLLSDQOLAESNALILA
QSETAKARARVQRIEHLATDDVDVAVTIDILSSVANDLRKKYLESSKLETEITRRIG		QSETAKARARVQRIEHLATDDVDVAVTIDILSSVANDLRKKYLESSKLETEITRRIG
SSHQAVRLNEMQREYRMLFQISIAQSYKSDVSEAKESLAESVAKAYVDISNS		SSHQAVRLNEMQREYRMLFQISIAQSYKSDVSEAKESLAESVAKAYVDISNS
ASSETQVRELQREAYETKNNYQTFLOREQAQQQSFVTEARVISKAMPYPSKP		ASSETQVRELQREAYETKNNYQTFLOREQAQQQSFVTEARVISKAMPYPSKP
NKPMILVFWAGAAAGGATIFREFRDFRTGDQVDVLGIFLGNTPLIPNQPTT		NKPMILVFWAGAAAGGATIFREFRDFRTGDQVDVLGIFLGNTPLIPNQPTT
EAQNGHPLGTRPTSVARVADHPLSSFAETLRSLRLAIDLGLIPAKTGARIVGWSL		EAQNGHPLGTRPTSVARVADHPLSSFAETLRSLRLAIDLGLIPAKTGARIVGWSL
PSEGKSTISINLAOLLAGOGARVLLDADIRNPGATRAMARHAAGLLEVLLEGRSVQ		PSEGKSTISINLAOLLAGOGARVLLDADIRNPGATRAMARHAAGLLEVLLEGRSVQ
DVLRDEKTYLAFPLTVVQKRVPHSSLELTSQMHKLLAEASAFYIIVDLPPLGPV		DVLRDEKTYLAFPLTVVQKRVPHSSLELTSQMHKLLAEASAFYIIVDLPPLGPV
VARAMAGRIDGFIFFTEWKTARRAVRVNEVHIRKKCLGVLINKVDTLEKLILYR		VARAMAGRIDGFIFFTEWKTARRAVRVNEVHIRKKCLGVLINKVDTLEKLILYR
AYGSSEYHSRYTRYHD"		AYGSSEYHSRYTRYHD"
6667..8190		6667..8190
/gene="Atu3557"		/gene="Atu3557"
6667..8190		6667..8190
/gene="Atu3557"		/gene="Atu3557"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"		/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1		/codon_start=1
/transl_table=11		/transl_table=11
/product="conserved hypothetical protein"		/product="conserved hypothetical protein"
/protein_id="AAL44369.1"		/protein_id="AAL44369.1"
/db_xref="GI:17741964"		/db_xref="GI:17741964"


```

CDS
<1..775
/gene="gmbd"
/codon_start=2
/transl_table=11
/product="ADP-L-glycero-D-manno-heptose epimerase"
/protein_id="AAL23753.1"
/db_xref="GI:16445332"
/translation="EFLIQIMAGEFEIEAIFHEGACSSTTWEGKYMMDNNYQYS
KEUHCYCLEREIEPLFVASSAATYGGRTSIESEYEQPLNVGYSKFLFDEYVRQIL
PEANSIVGFRYFNVYGPREGHGKGSMAFHLNTQLNNGESPKLFEKSGDFKRDVY
VGVADVNLFWENGSGIFNLGTGRAESFQAVADATLAYHKRGSIEYIPFPEKLRGR
YQAFTQADLTNLKAGYDKPKTKVAEGVTEYMAWLNRA"
785..1831
/gene="waaF"
785..1831
/gene="waaF"
/note="complements a Serratia enterica serovar Typhimurium
waaF mutant"
/codon_start=1
/transl_table=11
/product="heptosyltransferase II WaaF"
/protein_id="AAL23754.1"
/db_xref="GI:16445333"
/translation="MKILVVGPSWVGDMMSORLYRTLRLARYPQAIIDVMAWCRPL
LQRPVNGQALMPLGHGALGERRLGRALRANRYDRAYVLPNFKSALVFFPADL
PORTGWRNAGLGLNDVRLDKAAPPMLVQRYVALAYDKGRVORADDLPQPLNLPOL
RVSDEIAETTSAFNLTDSRPVYFCGPAEFGPAKRWPHYHYAALAAQLTESYQIVL
FGSAKDEAGEQIRAAALQDDARDFCLNLAGKTQLEQAVILIAACRAVVSNDGLMHA
AALNKLEALYGPSDFPTPLSDKARVIRLSGYHKVRKGDAEQGYHOSLIDIQPOQ
VLDALPLLVASEE"
1834..2799
/gene="waaC"
1834..2799
/gene="waaC"
/note="complements a Serratia enterica serovar Typhimurium
waaC mutant"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I WaaC"
/protein_id="AAL23755.1"
/db_xref="GI:16445334"
/translation="MOVLIIVKTSMDGLVHLTPALDQAIPIDFWVWVEGFSQI
PTWPAVDVPIVAIRWRKNWFGSDTROCRDFKALOEYRDVVIDAQLMKSAL
ITRVAGSKHGPKCKSARPPASFWYNVRHEIDKQHAVERTELFAKSLGYDKPGSY
GDVAARFLQSPVQDAGQVLFELHATRDCKHWPEONNREHLIALTADSLKIKLPMG
AEHEQRALAEGFPHVEVLPKLSQQVADVLGAGVVSVDTLGSLHTAALDKRPNL
TLFGPTDGLIGYGQNHSLISPEKSMATIDATAWAQLQKVIA"
2796..3962
/note="Orf4"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAL23756.1"
/db_xref="GI:16445335"
/translation="MKKLHIINLKGWGVVERFLQVINDTDCSNQVLCISGDIGBEI
RQGLPHQPVTFANRLNALPLKCPQFLKFLKWKIERANADVIVMDLVPGLAAP
RKGLVYDHGCSWRYPKNQKTRFFAMLDGVISASHASKRMELRFLPCPNHVIN
RIKTPADITNLSQPIRIGTASRLVSLKISVSLAMQELLRRGHVDVLEVAGK
PDRAFEALAAQLGDRVTFSGYQDDVAGFFNRTHYINSTPITEPFGLSMESLYFG
VPVIFPQDQPEAVKGVCGIGLVPSTIEOHRQLTDIEVDFPHEVIDPLTSLVSP
KLKSHIDCADAVETLVAGETYQLLSRNAQRYPAEHFYAQFAEFDLTRSFIV"
3991..5232
/gene="waaL"
3991..5232
/gene="waaL"
/note="similar to Escherichia coli K-12 WaaL"
/codon_start=1
/transl_table=11
/product="lipid A core: surface polymer ligase WaaL"
/protein_id="AAL23757.1"
/db_xref="GI:16445336"
/translation="MLKOKTPHAFSYLIYLGCAIAFCTIPFGSATGRNLFVYSSVIA
FIAVLPRYFNFNVNALLPALMFCVGMGTILMMHHFQKPGEYIYIYSYMTGKLIQ
LATAFILLIALNERLCVQRLIIVALTGTGAIINGYSIYQGLRIDIPRVEINFORATVA
AYLMTAIDLVMQAILMLRTRYRLVLYIAAFLLSFSAVLTGTMAMLVYPVAICLSL
LATKHLVSRKHALVASVPLLVAGLKFQPIERIQAFETDRMLMDQPQPTENSII
SLKSMOTLAWRTGSOAPWGSOAEORGEIRAIVAQOPRLSGVMPYINVHNLNLETFY
SLKSGTGALLLALYIGLTFSSFRPQNRALLSASLFVYGLSDVIFFTGTEGVFEC
LALIAGVLSVAKPPSQAQETA"
5229..6209
/note="Orf6; similar to TsrB (Wbck) protein from Versinia
enterocolitica"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAL23758.1"
/db_xref="GI:16445337"
/translation="MKTHQPTLLSVIPFYNNFAVIAALDSLTQISDEIEVVIID
DDYDASGALVSRYLAQRHQRVAFISOANGIAHARNVQLOHGAARVLTFLDGDLLLS
DDYLAIRPLLIADVEDLIDFNQKTEQPKPENRATOPAAVDFSLGSLQPLFA
RSMHLWSRYVKTLLAGETFEDEGRYEDVIFPFQYFKRRIAHLDCELYFTRDNSQ
GITRNKPKDIEDLLFAMKMLRFVAQHPQDEPLRLAALMLANCSEVSKMSKAVYG
YHYAPATIQILRRATVCNSDVPCCKVKQMYARVDVDFLSKLRGKPR"
complement(6246..7340)
/note="Orf7; LPS from a non polar orf7 mutant has reduced
levels of D-D-heptose"
/codon_start=1
/transl_table=11
/product="heptosyltransferase"
/protein_id="AAL23759.1"
/db_xref="GI:16445338"
/translation="MSKKWKIKVADMFRLRYTAKYGFQDKATFDCAQOQDNIYIST
TALGDMFNTPALRAMERYPNAHILVHVKKRELVENGRYDRVYVWSKIKTIRP
LKLSHQHPDLALILSHLPIYDIISAVMAGAKYIINDNTSSGICGLEPMTNIFYH
YGHFVRKKLELISILGRSDPTIMEIPVDYQGVKSDHPTFQGLGASERSCWPVS
HYARLADDELVAKHPNVRILVILGAPLDQAKQFMSLTGPQTCVQNDQIGKTQGLL
TLDSFHLVLTGDTGPLHLAVALKVKITISLFTVADSRSTGPIQDPELLRIQVSRTEY
TMLEAESPMGVIOQRIYDELVAHVPLAR"
8516..9598
/gene="waaQ"
8516..9598
/gene="waaQ"
/note="complements a non polar Klebsiella pneumoniae waaQ
mutant; similar to WaaQ proteins from the
Enterobacteriaceae"
/codon_start=1
/transl_table=11
/product="putative heptosyltransferase III WaaQ"
/protein_id="AAL23760.1"
/db_xref="GI:16445339"
/translation="MMDAPALTSAQSVQRILLIKLRHHGDMLLVTPVSISSLRQNYPO
ARIDVLYQTEMLASNPESLTIFAIDRWKKQGAHGLHGLALRLRKAQRIYDILV
VNLADQWRSALITRLTGARIRLGLDFPKRRGFLWRHCHTQLVPVDNNAHLHTVQNL
LLAPNLNLPALNEHTMSYDPDQWQTCQLLQKQGVAGSYIVVQTSWFFKWCSEKM
AATLTALQADGHOLVITSGPDAREKAMVERILALCPQGVISLAGOLTLRLAALIDH
AKLFIGVDSVPMHMAALQPCVALFGPSKLTFRWPQATGAVIAGDFGELPDPDAI
DTGTDERYLDLIPTDVIAAARSTLA"
9595..10722
/note="similar to Klebsiella pneumoniae C3 putative
glycosyltransferase and from salmonella enterica
serovar typhimurium, and Escherichia coli strains K-12,
R2, R3 and R4; Orf9; LPS from double non-polar orf9-orf10
mutant lacks GalA, has reduced Glc levels and contains
O-antigen"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAD28801.1"
/db_xref="GI:4753135"
/translation="MKAFLLAIYRRKYRYPDGAERFVSRAKALEQOQDLNLVITREW
QGANPNWHILCNPLKLRISRGFAVARALWOKEREDLVOSHERIPCCDIYRAG
DGHVRWLLQRLARLLPEWRKWLFSNRYHRYVMCAERAMTAAPELXACIENAGWIKRE
IYHDFGVPAKTIYVNAIDNQKFPFPADEQRRLREYOIIPQAHCLIFVSGFKERK
GLAAATRAVAATDHLVGVCKDAEKRYRALQSLGCGDRIFHEMGYKQKOTLPFYQAD
ALLPLYDFPNVILEAMSCGLPVITSTTCGGAETFTPCNGCFVTDALDVPALTEAI
RALPRQALGSSMGAAKLRIWTATPAHLSEQLISLYNRLID"

```

CDS

Query Match 4.3%; Score 50.6; DB 1; Length 20693;
 Best Local Similarity 50.2%; Pred. No. 0.00034;
 Matches 125; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 382 CATTCCTCTCAATGACATCAGGATCCGCCCGCTCTCACTGGCGATACAGCGGCGG 441
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 11667 CTTTTCAGAAATTCGCGGAGCGCGCGGAGCGGCGGATGACACCGCGGTCGCCG 11608
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 442 AGACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAGGCGATGACCA 501
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 11607 ACAACAGGATTCACGAGCAGATTCGCCAACCTTCGCTATCGGAGTGCAGCACCA 11548
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 502 CACTGGCAATCCGCTAGACCGCTAACGCTGGGAAAAGGCGATGCTTCAATTAACATCTC 561
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 11547 GCGAGGATGCGGATACAGGATAGGATGCGTGGAAACCGAGGACGACGCGCT 11488
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 562 CGCTCATTCAGGCTGCTCTGCTGAGCGCAGAGCTGCTTCTGATTTCTACGCGCGG 621
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 11487 CTGCGATACCAAGTCTGCCGCCAGAGCTTGACCTCGCGGCGGCGTATCCCCCG 11428
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 622 CGCCACCA 630
 ||||| |||||

DB 11427 TGCGGATCA 11419

RESULT 29
 LOCUS AE013034/c
 DEFINITION Thermoanaerobacter tengcongensis strain MB4T, section 61 of 244 of the complete genome.
 ACCESSION AE013034 AE008691
 VERSION AE013034.1 GI:20515643
 KEYWORDS
 SOURCE Thermoanaerobacter tengcongensis.
 ORGANISM Thermoanaerobacter tengcongensis.
 Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
 1 (bases 1 to 8594)
 Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Yu, J., and Yang, H.
 A Complete Sequence of the T. tengcongensis Genome
 Genome Res. 12 (5), 689-700 (2002)
 2 (bases 1 to 8594)
 Bao, Q., Xu, Z., Hu, S., Dong, W., Chen, Y., Wang, J., Yu, J., and Yang, H.
 Direct Submission
 Submitted (07-OCT-2001) Beijing Genomics Institute/Genomics and Bioinformatics Center, Institute of Genetics and Development, Chinese Academy of Sciences, Beijing Airport Industrial Zone B6, Beijing 101300, China
 3 (bases 1 to 8594)
 Li, W., Xuan, Z., Yang, J., Ling, L., and Chen, R.
 Direct Submission
 Submitted (07-OCT-2001) Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
 4 (bases 1 to 8594)
 Tian, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., and Tan, H.
 Direct Submission
 Submitted (07-OCT-2001) Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, China
 Location/Qualifiers
 1. 8594
 /organism="Thermoanaerobacter tengcongensis"
 /strain="MB4T"
 /db_xref="taxon:119072"
 102..1289
 /gene="RfaG"
 /note="TTE0655"
 102..106
 /gene="RfaG"

FEATURES

Source
 1. 8594
 /organism="Thermoanaerobacter tengcongensis"
 /strain="MB4T"
 /db_xref="taxon:119072"
 102..1289
 /gene="RfaG"
 /note="TTE0655"
 102..106
 /gene="RfaG"

gene

RBS

117..1289
 /gene="RfaG"
 /note="Best Blastp hit = gi|2129096|pir||F64500 probable hexosyltransferase (EC 2.4.1.-) M1607 - Methanococcus jannaschii gi|1500505|gb|AAB99629.1|'(U67601) LPS biosynthesis protein, putative [Methanococcus jannaschii]', score 100, E-value 3.00E-20"
 /codon_start=1
 /transl_table=11
 /product="predicted glycosyltransferases"
 /protein_id="AAM23922.1"
 /db_xref="GI:20515644"
 /translation="MRDNRIQDVALLISSNPYITNLGQLHLLLEKRLNNVNVYTKYTFSSFEKAKLVFPYFKFMSYKDRVITRTKLEKFFKQHFQVFNADVI SGACNSKIKVILTHGYTRAINYGFSGEDIPDFERGLDEKRAQRIITVD SRKGVVSESPFPEKVDVIFNAVDIDFKPISLEKKQIREELGWGEEDVVFPIR RVKNGVYAKAANILKEKDKIFAFIGLPEKEIILNTHKNNKVLKGLGPHDEI VRYKACDVLIPSTSDGVEATSLSMLEMSCGKIYVCTPIGCKMELIKHGVNGFF VQKSEEAIAIYIIKEKEDFYKLDISIRQEARKIIEKNHSYIVHANKFIEVYKAIK"
 705..1217
 /gene="RfaG"
 /note="Pfam match to entry glycosyltransf.1, Glycosyl transferases group 1, score 123.5, E-value 3.90E-33"
 1350..2742
 /gene="TTE0656"
 1350..1354
 /gene="TTE0656"
 1369..2742
 /gene="TTE0656"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAM23923.1"
 /db_xref="GI:20515645"
 /translation="MFTLVILGLMIVETEIMWRKIKFDLFRGNIVYELIYFL ARSYIILDFKVDSLPSSMWFLINYNKDAYSWFLASVISVGFCTVGVYGLGISF RNIGCSINFKLNKISDSIWLSSIIIVLFGFSGLYIKRSAPFDIMISGLLRSC IKVGVETSTFTFLSVTFSSYVLLGLSLNFKFSKFLFLYGLFGISGLSLV LILKAGRLHNLVLLVFLTYSTLEKIRFLFLSVLSLSLILYKGYWFGIRFG KPIYNSLOQIFSNFAEFPYLSLINLCNPKYTSRYFLDFPLGLIRPLGLISK IIDINDPLTSQENTLIIIGTEKGIIPVDILGAGVFSANIFGVTFMFCGLMLAL FERIFSNVSNPVLRSILKIAVILYFSTMVLYADPAILLMDGVYLIFPIVFLMLYLL NEFKYSP"
 2751..3904
 /gene="RfaG"
 /note="TTE0658"
 2751..2755
 /gene="RfaG"
 2762..3904
 /gene="RfaG"
 /note="Best Blastp hit = gi|7433884|pir||E72354 probable hexosyltransferase (EC 2.4.1.-) TM0622 - Thermotoga maritima (strain MSB8) gi|4981142|gb|AAD35706.1|AE001736_4 '(AE001736) lipopolysaccharide biosynthesis protein, putative [Thermotoga maritima]', score 263, E-value 4.00E-69"
 /codon_start=1
 /transl_table=11
 /product="predicted glycosyltransferases"
 /protein_id="AAM23924.1"
 /db_xref="GI:20515646"
 /translation="MNRKIVQLIIPNFGVGAERLTVMKYLDEKKEVEYRAISMFDS LNTELEKLENIPIYVYLGKKGDFPRMFRIDIKISFPHIVTHRYVLRVALPS LLLHKVPKVHVHNAEKVDKGLVKAFISFGVTPISIRLSVLSLSLTVGVKN IPLILNGIPVEYQKANIENREWEKGEFQEDFLVNIARLAPKNALLTEAFAGK PARHNSKLIIVGDGEERLEITKLHLEKVEYELGIRWDIPILNNSDVFVLSDD WEGNLSKNEAAKAEKFDIATSVGGVPELIQNNITGILVPKNVNAFKRAMLMLIEN DICQKLGEKAKAEKFEKFDISVWVKYKLEKYESLQFLKKGASLL"
 3296..3802
 /gene="RfaG"
 /note="Pfam match to entry Glycosyltransf.1, Glycosyl transferases group 1, score 199.9, E-value 4.00E-56"
 3886..4839

misc_feature

gene

RBS

CDS

gene

RBS

CDS

misc_feature

gene

```
/gene="TTE0659"
3886..3890
RBS
/gene="TTE0659"
3901..4839
CDS
/note="Best Blastp hit = gi|13442978|gb|AAK26246.1|
(AF330049) D-glucuronyl C5-epimerase [Mus musculus], score
96.3, E-value 5.00E-19"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM23925.1"
/db_xref="GI:20515647"
/translation="MYVRLSKYLRLLGKSLAHVQNGIGKIYSKEGKGYNDLTL
KYNHGETKRGLPNMVNDGKGIYPTTVIQGLGNYDILITKQDITFEQWKSVM
WVNEDEGGWDFERIGSKPKFYKSAMTQEGASLLIRAFQESDEDFVDRAKAI
DMLLPVEGGLARYYNDKLFEEVYESKPLNGWIFAVFLGYDIFRATSDGRKE
ALQRTLDLTKELYKYDTGWSYIDQGNLSPYHKLHIALLEVLVELFNTIEFTV
KKEWSYLSNKKRTRAFIVKAYOKIKEPGEVLIK"
4840..6072
/gene="RfaG6"
/note="RfaG6"
/gene="TTE0660"
4840..4844
RBS
/gene="RfaG6"
4852..6072
CDS
/gene="RfaG6"
/note="Best Blastp hit = gi|7470136|pir|I577339
hypothetical protein sll1723 - Synechocystis sp. (strain
PCC 6803) gi|1652521|dbj|BAAL7442.1| (D90906) hypothetical
protein [Synechocystis sp.], score 117, E-value 3.00E-25"
/codon_start=1
/transl_table=11
/product="predicted glycosyltransferases"
/protein_id="AAM23926.1"
/db_xref="GI:20515648"
/translation="MKGMIAYITSTPYKGGQFQILPELLEVIKKGHDVVIPVRPE
KETALGDEPKVAKFMHPIPLSLILKSLFIFKHLRVVVEILFKILRHSGSLGKI
LKNLLIIPKGLVSEILKRRKINHIAHWASPTSLAYASELTGIPWFAHRWDIA
ENMLREKARSNFIRVDKNGYDEILSVKEYRDKLIYLHVGKRVYVNFKLDLKKKA
NRALLIYANFVYKGGHLYIOAIKKVITQNCYCYFFGKGLLEDLKLTVTQLN
LEDRIFFQGLPHELIYRSGKIDCVILPSIVTDGKEGIPVSLMEAMAHKIPVI
STNGGIPELLEGGAGIIVEQKNSDELAKAIKMLINDEKLRELGEGKEFKIEKFN
SKIVEELLYLMGN"
5470..6006
/misc_feature
/gene="RfaG6"
/note="Rfam match to entry Glycos_transf_1, Glycosyl
transferases group 1, score 154.5, E-value 1.80E-42"
6085..7190
/gene="WecB2"
/note="WecB2"
/gene="TTE0661"
6085..6089
RBS
/gene="WecB2"
6099..7190
CDS
/gene="WecB2"
/note="Best Blastp hit = gi|3451512|emb|CAA07668.1|
(AJ007747) putative UDP-N-acetylglucosamine 2-epimerase
[Bordetella bronchiseptica], score 381, E-value 1.00E-105"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine 2-epimerase"
/protein_id="AAM23927.1"
/db_xref="GI:20515649"
/translation="MKVITIGTRPQIFKAATVSRKLFNDMKKEIIVHTGQHDYDM
SDIFFELIPQPDYLYIGSGSHGVTAMLEKIEDVLIEKPDVLYIGDTNSTLA
GTIAASKLQIPVAHVAGLSRFRKMPPEINRIVADVSDLLFAPTKVAVNNLNEGI
ADRIYVGVDMVDAIFYGKAGQSKILETLDKPKVEYILSTIRAYNTDIPERLK
TIFNALCKIGKQITVWPLHPPTKALLREGLYKVESEKLIIEPVYKLDVMDKLENA
KLITDSGGVQKREAFYKVLVTLRETEWVELVGLGNVLPVPPDCEKFTIESIRKVL
NAPHGLDAPPIYDGGKAAEKIVKVLRYLL"
6228..7175
/misc_feature
/gene="WecB2"
/note="Pfam match to entry Epimerase_2,
UDP-N-acetylglucosamine 2-epimerase, score 512.8, E-value
```

```
2..50E-150"
7215..8462
/gene="RfaG7"
/note="TTE0662"
7215..7219
RBS
/gene="RfaG7"
7227..8462
CDS
/gene="RfaG7"
/note="Best Blastp hit = gi|3451511|emb|CAA07667.1|
Query Match 4.2%; Score 49.6; DB 1: Length 8594;
Best Local Similarity 48.9%; Pred. No. 0.0006;
Matches 133; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
QY 232 TCTGTCCGCTTGTGTTAATATCAACCGGTACTCAATATCTTCTGTGGCGCTGGCTGCCA 291
Db 3839 TTTTACCATTACACTAATCAAACTCTTCTCAGCAACTTCTTAGCCTTTTCTCCTA 3780
QY 292 TCATCCGGAGCGCTTCGGTCCGGATAAAAAATCGGCAGTGGCCGGTCCATGCACACA 351
Db 3779 ACTTCTGACATAAATCTTTGTTCTCAATTAACATCAGCATGCTTTTAGAAAAAGCGTTTA 3720
QY 352 CATCCCCCAGCGGTAAACAGCGTCCCTGTCTCATTCTTGAATGACATCAGGGATCCGC 411
Db 3719 CATCTTTGGGGGAACAAGAAATTCCTGTAATATATTTCTGTATACTCTGGAACTCCTC 3660
QY 412 CGCTCTCAGTGGCGATAACGGGCGAGGAGACTGACGCTTCAGCCAGTACCATACCAA 471
Db 3659 CTACTGACGTTGCTATCACCGGTTTCCAGCAGCATGCTTCCATAACACTTAGTGGAT 3600
QY 472 ACCTTCATTTTCCGAAGGCATGACCACACA 503
Db 3599 TGCCTTCCCAATCTGAAGATAAAACAACACA 3568
RESULT 30
AC005674/C
LOCUS Homo sapiens chromosome 4 clone RP11-448G15, complete sequence.
DEFINITION AC005674
ACCESSION AC005674
VERSION AC005674.11 GI:13129492
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 199450)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 199450)
AUTHORS Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-1998) Department of Genetics, Stanford Human
Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 3 (bases 1 to 199450)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Feb 27, 2001 this sequence version replaced gi:5757495.
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.2.
STS Content:
SHGC-51391 G34042
SHGC-69010 G41975
SHGC-50921 G33945
SHGC-25155 G31610
SHGC-51591 G34129
SHGC4-1311 G33196.
Location/Qualifiers
FEATURES
```

```

source          1. .199450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-448G15"
BASE COUNT     54835 a 46671 c 45223 g 52721 t
ORIGIN
Query Match      4.2%; Score 49.6; DB 9; Length 199450;
Best Local Similarity 52.4%; Pred. No. 0.0012;
Matches 109; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 74 TCTGTCGGGATTTAAATGATTCACCGTCCCGTATTTAAACAATGTGATAAATPAC 133
Db 112372 TCTGTTGGAGCTAGTGCCACTGACTCTTTGGGTGTACATCAGAATCACCTAGTC 112313
Qy 134 CGTTACCGGAAACCGCTGACAAATTCGGGCTGAAAGAGGATCCGCGTATCTGTT 193
Db 112312 CCTGTCCCAACACACTGAATCAGAATGGCAGCGATGAGACTCTAGTGAATGTGAAGTT 112253
Qy 194 GCATTTCCCTTGTAGCTGACTAGCAGACACAAATGATGTGCGGTTCTGTGTAATATC 253
Db 112252 GTATTTCCCTTGTGGTCTCATTTCTCTGTGACTAATGATGTGAGGCAATCTGTGATTAGC 112193
Qy 254 AAACCGGTACTCAATATCTTCTGTGGCG 281
Db 112192 ATACCACTCATTTGATTTCTCTGGAG 112165

RESULT 31
AX433280/c      1135 bp DNA linear PAT 28-JUN-2002
LOCUS           Sequence 1695 from Patent WO0229113.
DEFINITION      AX433280
ACCESSION       AX433280
VERSION         AX433280.1 GI:21658084
KEYWORDS        Bacillus licheniformis.
SOURCE          Bacillus licheniformis.
ORGANISM        Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE       1
AUTHORS         Berka, R. and Clausen, I. G.
TITLE           Methods for monitoring multiple gene expression
JOURNAL         Patent: WO 0229113-A 1695 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES        Location/Qualifiers
source          1. .1135
/organism="Bacillus licheniformis"
/db_xref="taxon:1402"
BASE COUNT     327 a 251 c 278 g 279 t
ORIGIN
Query Match      4.1%; Score 49; DB 6; Length 1135;
Best Local Similarity 50.2%; Pred. No. 0.0006;
Matches 121; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 351 ACATCCCCCAGGGTAAACAGGTCCTCGTGCATATCTTCTGAATGATCATCAGGATCCCG 410
Db 977 ATATCTCCGACATCGACCAAGTAAACCGGATACCTCCGTTTTTGTATCATCCTCGGGAATCCCG 918
Qy 411 CCCGTCCTCACTGGGATAACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCA 470
Db 917 CCGATATTGTGCGGATGACGGAACCGCCGAGCCCATCGCTTCAAGCAGCACTAAACCG 858
Qy 471 AACGCTTCATTTCCGAAGGATGACCAACACACTGCGCAATCCGGTAGACCGGTAAACGCT 530
Db 857 AAGCTTTCTCTCGGACAAAGACAGCTTTAAATCGTGTATGCGGTAAACCTCTTCTACG 798
Qy 531 GGGAAAGGGCACTGCGCATTAACATATCCCGTCAATTCGCCAGGTGTTCTGTGCTGTA 590
Db 797 CTGTCTTTGTTGCCAAGAAAGAAACGTCGTCTTTTAATCCGAGCTGCTGTGACAAGCTGG 738
Qy 591 C 591

source          1. .783
/organism="Bacillus clausii"
/db_xref="taxon:79880"
BASE COUNT     225 a 149 c 199 g 208 t
ORIGIN
Query Match      4.1%; Score 48.6; DB 6; Length 783;
Best Local Similarity 49.8%; Pred. No. 0.00075;
Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 354 TCCCCAGGGTAAACAGGTCCTCGTGCATATCTTCTGAATGATCATCAGGATCCCGCCC 413
Db 631 TTCCCAAGGGGCATATAATAAACAGTATTCTGCTTCAACAACATTCGCGGATGCCACCA 572
Qy 414 GTCTCACTGGGATAACGGGACGCGGAGACTGACGCTTCCAGCCAGTACCATACCAAC 473
Db 571 ATTTTGTGCGATCAGAGTACCGGATCGCATCGCTTCAAGCGCCACTAAACCAAG 512
Qy 474 GTTTCATTTTCGAAGGATGACCAACACACTCCGCTCAATTCAGGTTAGACCGGTAAACGCTGG 533
Db 511 CTTTCTTTTTCGTTAAACAACAGCATTAATCGCTCATAGACACAATTCAGCAACATGC 452
Qy 534 AAAAGGGCACTGCCATTAACACACTCCGCTCAATTCAGGTTAGACCGGTAAACGCTGG 593
Db 451 TTTTGTGTCCTAGCATAGGACATCGTTTGTGAAGCCGTTTCTTCCACAACATCGCTC 392
Qy 594 AGACGTG 600
Db 391 GCAATTG 385

RESULT 33
AP000007/c      253505 bp DNA linear BCT 06-APR-2000
LOCUS           Pyrococcus horikoshii OT3 genomic DNA, 1485001-1738505 nt. position (7/7).
DEFINITION      AP000007
ACCESSION       AP000007 AB009465 AB009521 AB009522 AB009523 AB009524
AB009525 AB009526 AB009527 AB009528 AB009529 AB009530 AB009531
BA000001
AP000007.1 GI:3236134
KEYWORDS        Pyrococcus horikoshii (strain:OT3) DNA.
SOURCE          Pyrococcus horikoshii
ORGANISM        Archaea; Euryarchaeota; Thermococci; Thermococcales;
Thermococcaceae; Pyrococcus.
REFERENCE       1
AUTHORS         Kwarabiyasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y.,
Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y.,
Sakai, M., Ogura, K., Otuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y.,
Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushiida, N.,
Oguchi, A., Aoki, K., Nakamura, Y., Robb, T. F., Horikoshi, K.,
Masuchi, Y., Shizuya, H. and Kikuchi, H.
TITLE           Complete sequence and gene organization of the genome of a
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3

```



```

/protein_id="AAL82012.1"
/db_xref="GI:18894079"
/translation="MFRPLRLVSIILVGLVFLVQVYFVGOAILGGRRSLREATG
VGDYPIYLVIGVLMWVSVPMSEASVWGVRLELQRTGFESNVMAPIGILKIMGLAIS
WMMSIIMLLFALGVWIFDIKSAIYYAKAIPAIALAFVFLGFIYFAGLVMLK
NIGPLANIEFVILFSLGVFFSLTPEIVRKISWIPIPLTHATSAVRKIFIGLSYSSV
VNEITAMLILLPVYWAISLAVFKWAEKVRMMGYGY"
complement(9079..10035)
/gene="Pfl1889"
/notes="Function Code: 15.2 Other Categories: Drug and
analog sensitivity"
/codon_start=1
/transl_table=11
/product="daunorubicin resistance ATP-binding protein
drrA"
/protein_id="AAL82013.1"

```

```

Query Match      4.0%  Score 47.4;  DB 1;  Length 10991;
Best Local Similarity 60.5%;  Pred. No. 0.0033;
Matches 78;  Conservative 0;  Mismatches 51;  Indels 0;  Gaps 0;

QY 350 CACATCCCGGAGGTAACAGGTCCTCGTCACATTCCTGTAATGACATCAGGGATCC 409
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6404 CTCATCCCGAGGGGCACTAGAGGCCAGCTTCCTCTTTATGATTCAGGTATCC 6345

QY 410 GCCCGTCTCAGCGGCAATACAGGGCACGCCGAGACTGACGCTGACCGATCATTACC 469
      || || || || || || || || || || || || || || || || || || || ||
Db 6344 ACCAAGCTCTAGCTACAACTGGGACTCCTGATGTCATTCGCTCGACCACTATACC 6285

QY 470 AAACGCTTC 478
      ||||| |||||
Db 6284 AAATGCTTC 6276

```

```

RESULT 37
AE012967
LOCUS      10148 bp      DNA      linear      BCT 01-JUL-2002
DEFINITION Chlorobium tepidum TLS section 188 of 194 of the complete genome.
ACCESSION AE012967 AB006470
VERSION    AE012967.1 GI:21648252
KEYWORDS
SOURCE     Chlorobium tepidum TLS.
            Chlorobium tepidum TLS.
            Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
            Chlorobium.

```

```

REFERENCE
AUTHORS    Eisen,J.A., Nelson,K.E., Paulsen,I.T., Heidelberg,J.F., Wu,M.,
            Dodson,R.J., Deboy,R., Gwinn,M.L., Nelson,W.C., Haft,D.H.,
            Hickey,E.K., Peterson,J.D., Durkin,A.S., Kolonay,J.L., Yang,F.,
            Holt,I., Umayam,L.A., Mason,T., Brenner,M., Shea,T.P., Parksey,D.,
            Nierman,W.C., Feldblyum,T.V., Hansen,C.L., Craven,M.B., Radune,D.,
            Vamathevan,J., Khouri,H., White,O., Gruber,T.M., Ketchum,K.A.,
            Venter,J.C., Tettelin,H., Bryant,D.A. and Fraser,C.M.
            The complete genome sequence of Chlorobium tepidum TLS, a
            photosynthetic, anaerobic, green-sulfur bacterium
            Proc. Natl. Acad. Sci. U.S.A. 99 (14), 9509-9514 (2002)
            12093901

```

```

2 (bases 1 to 10148)
Eisen,J.A., Nelson,K.E., Paulsen,I.T., Heidelberg,J.F., Wu,M.,
Dodson,R.J., Deboy,R., Gwinn,M.L., Nelson,W.C., Haft,D.H.,
Hickey,E.K., Peterson,J.D., Durkin,A.S., Kolonay,J.L., Yang,F.,
Holt,I., Umayam,L.A., Mason,T., Brenner,M., Shea,T.P., Parksey,D.,
Nierman,W.C., Feldblyum,T.V., Hansen,C.L., Craven,M.B., Radune,D.,
Vamathevan,J., Khouri,H., White,O., Gruber,T.M., Ketchum,K.A.,
Venter,J.C., Tettelin,H., Bryant,D.A. and Fraser,C.M.
Direct Submission
Submitted (30-Apr-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
unpublished
FEATURES
            Location/Qualifiers
            1..10148

```

```

/organism="Chlorobium tepidum TLS"
/strain="TLS"
/db_xref="taxon:194439"
complement(242..1099)
/gene="Ct2209"
complement(242..1099)
/gene="Ct2209"
/notes="identified by match to PFAM protein family HMM
PF00535"
/codon_start=1
/transl_table=11
/product="glycosyl transferase"
/protein_id="AA073425.1"
/db_xref="GI:21648253"
/translation="MNGKTAIVIVLGNAGADTLACIASLAKVRQPAFTVILADNGST
DETGLVROAFPEVEILEGRNIGFAAGNNAAFRSRSGRGTDFRVFLNNDTVVDPGFL
QPLIDELQRPWYIAAPKILYIMDDPGRIWAGVLESATGLIHTGIOPDGPFRFTFP
EPVYATGCCLAMRCRVFEVGGFDERFMYGDDVLSMKVREGLIVYQAPASRIWH
RVSSSGGEMNLGKQLKSKAAMMLFAKHGMIGGLVLYPLLLFRALLGLLRQPFWM
TSSQREEA"
complement(1096..2121)
/gene="Ct2210"
complement(1096..2121)
/gene="Ct2210"
/notes="identified by match to PFAM protein family HMM
PF00534"
/codon_start=1
/transl_table=11
/product="glycosyl transferase"
/protein_id="AA073426.1"
/db_xref="GI:21648254"
/translation="MFRRLYPRWLNAGRTATPEGSDVPASPTFDLMNPLSWITAA
RKLNRKIPDVLIAYWSGLAPICALMRRAAGLPTFVLLHNETSHEPISGESMLKRL
VSSDGFITISOTVESELAFAPAAKTLRLFLHLYBROTPGPSKAARRSIGLPEDAP
VLFFGVREYKGLDTLFEAMALVLRSSARLVAVAGEFLGSSRFREARRRIGDGA
VEFREGYPAGEVATLMAADAVLVPYSATQSGIVPLALGHGVPVACDTGGIGNOV
EHGRTGLVREGEALADGILDFFRERELPLEEGISDFRRNSRREASQAATLE
TCSRSA"
complement(2273..2539)
/gene="ptsH"
/notes="Ct2211"
complement(2273..2539)
/gene="ptsH"
/notes="identified by match to PFAM protein family HMM
PF00381"
/codon_start=1
/transl_table=11
/product="phosphocarrier protein HPr"
/protein_id="AA073427.1"
/db_xref="GI:21648255"
/translation="MIVQEVIRNSAGLHTRPAAAVVKLASRFSDFEIMDGLIENA
KSIIGVMSLAAPKGRMVLKGEDEAAAKHLIEFEEQFGGEA"
complement(2604..3071)
/gene="Ct2212"
complement(2604..3071)
/gene="Ct2212"
/notes="identified by match to PFAM protein family HMM
PF00583"
/codon_start=1
/transl_table=11
/product="acetyltransferase, GNAT family"
/protein_id="AA073428.1"
/db_xref="GI:21648256"
/translation="MPAADTCVRNARLADASAIRTEGYAGEGIMLKRSVENIIEHI
RQFFVADYKGOVICCAIAFYTVKLAEIRSLAVLEFRNKGIGRLIVERKAEVLSSEG
VNEVFVLTLSNGFFKRMGYKEIEYFPQKIWRDCTNCPKRMACDEIATAVKTL"
complement(3058..4065)
/gene="obg"
/notes="Ct2213"
complement(3058..4065)
/gene="obg"
/notes="identified by match to PFAM protein family HMM
PF01018"

```

```
/codon_start=1
/transl_table=11
/product="GTP-binding protein Obg"
/protein_id="AAW73429.1"
/db_xref="GI:21648257"
/translation="MKFVDSAKISVKGAGDGGRCVSRREKVPKGGDGGDGRGGH
VYLRANKQTLTLDYFKRSYTAGRGHGLKARKSGDKGKDVIIIGVPGTVVVRNVEVG
ELVCDMEQOETWIAKGGGRCGNGOHFATRAOPRPAQCEGEEYELEWELKMA
DVLGVPNAGSTLISLSARPKIADYPTTLVPLNGLIVRYEDYKSFVADIPGII
EGAEGRGIGQIFRIERTKILLIMVPSNEDIAEAYATLKELEKEDPSLSAPRL
VITRMIDIAPEFTMLEKGVKVLTAISSVAGGLKALKDELWROVSLQNSPSEHAG
S"
4135..4422
/gene="xseB"
/note="CT2214"
4135..4422
/gene="xseB"
/note="identified by match to PFAM protein family HMM
PF02609"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease, small subunit"
/protein_id="AAW73430.1"
/db_xref="GI:21648258"
/translation="MPASKSRTSAYPTIEELIQRLEETRNIEPNPTDGLNSIALYE
EGSLAPECRKRLKETRKLETPAETARPAPKPNAPESPRMNDLFGTES"
complement(4437..5864)
/gene="gatB"
/note="CT2215"
complement(4437..5864)
/gene="gatB"
/note="identified by match to PFAM protein family HMM
PF01162"
/codon_start=1
/transl_table=11
/product="glutamyl-tRNA(Gln) amidotransferase subunit B"
/protein_id="AAW73431.1"
/db_xref="GI:21648259"
/translation="MNYEIVVGLVHVCQLNTESKAFCCGSAKFGKPANTNVCVCLAL
PGALPVLRNARVVEDAVKLGATNCTIARHSILARKNKFYDPLPKGYISQYEPIGSE
GVTHIDLEEGKDVLRVRIHIEDACKSIHDIGDITYIDNRCGVPLLEIVSYPMRT
PKRSAYLQKLRIVYLGISDGNMSEGLRCDANYSVRPGATEYGTETIKMNSF
RNVRLEYDAKHRIEVEGGGTIVQETRLWDADKLETRMSRGRKRAHDIYRFPDPL
VPLVDDGMTRHOEEPEPEPDRAARFSEFGIPAYDAGVITVDRELADYFESTVKV
SGDAKSNVVMGVEMTLKEKYLIDHKFAISPERLGLIKINAGIAINTAKQVE
IMQODEATAIIVEREGLAQSVDKGAIEAIREILEANOKLEQYRSGKTLQFGFVG
CQKMKGKANPKMNDILRSLDA"
complement(5910..6491)
/gene="CT2216"
complement(5910..6491)
/gene="CT2216"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAW73432.1"
/db_xref="GI:21648260"
/translation="MILLSNEKVMAMKRVHLMRALIALAATVLMFGTARAGWDPAD
EVNARTIGMFRKADPSLRFKSGYAVFSDIYKGFMIYGGHGKGVVFDGGRPI
GHAVTFNVPOLGGSFSEILIFFKDRALADFTKGNYSQAQFAVAVRAGMATNT
DYSNGVAFAMPNAGLMGLTVGAQKFTFRQYR"
6475..6603
/gene="CT2217"
6475..6603
/gene="CT2217"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAW73433.1"
/db_xref="GI:21648261"
/translation="MDNKIRPAKSNATPAYNVRFSSLKTRIASERKKPKNQLVGR"
6681..6824
```

```
/gene="CT2218"
6681..6824
/gene="CT2218"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAW73434.1"
/db_xref="GI:21648262"
/translation="MQGHAISIQPRCVITPMLYAFELATQCQPAYSAFDSIYTYFP
WRY"
complement(6837..7355)
/gene="CT2219"
complement(6837..7355)
/gene="CT2219"
/note="identified by match to PFAM protein family HMM
PF01661"
/codon_start=1
/transl_table=11

Query Match      4.0%; Score 47.2; DB 1; Length 10148;
Best Local Similarity 50.4%; Pred. No. 0.0038;
Matches 115; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 281 GCTGCTGCCATCATCCGGAAGCGTTCGGTCGGGATAAAAAATCGCCAGTGCCTGGT 340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1183 GATGCCCTCTCCAGCGGCGGCTCTCGCTCGCGGAAAGTCGAGAATCCGTCAGC 1242
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 341 CCATGCAGACATCCCCACGGGTAACAGCGTCCCTGCATCTTCTTGAATGACATC 400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1243 CAGTCCCTCGGGGCCCTCTCTCGCGACGAGCCGCGGTACGTCGCTCGACCTGATT 1302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 401 AGGGATCCGCCCGCTCTCACTGGCGATAACGGGCACGCCGAGAGTACGCTTCAGCCAG 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1303 GCCGAGCGCGCGGTGTCGAGCGCATCACCGGCACGCCGCGGAGCGGAGTGCGCAC 1362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 461 TACCATACCAACCGTTTCATTTTCGAAGGATGACACACACACTGGC 508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1363 GATGCCCGACTCGTGGCGGAGCGGTAGGCGAGCACCACCGCATCGGC 1410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 38
LOCUS AX437696 783 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 6111 from Patent WO0229113.
ACCESSION AX437696
VERSION AX437696.1 GI:21662504
KEYWORDS Bacillus clausii.
SOURCE Bacillus clausii.
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Berka,R. and Clausen,I.G.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 6111 11-APR-2002;
FEATURES
    source
        1..783
        /organism="Bacillus clausii"
        /db_xref="taxon:79880"
BASE COUNT 224 a 150 c 204 g 203 t 2 others
ORIGIN

Query Match      3.9%; Score 46.2; DB 6; Length 783;
Best Local Similarity 59.5%; Pred. No. 0.0046;
Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 354 TCCCCACGGGTAAACAGCGTCCCTGTCACATTTCTGATCATCATCAGGATCCGCC 413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 633 TTCCCAAGGGGCGATAAATAACCAAGTTACTTCGTCTTCAACACTTCCGGGATGCCACCA 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 414 GTCTCACTCGGATACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATACCAAC 473
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db	573	ATTTTGTGCGATCACAGTACCCGCGATCGCATCGCTTCAAGCGCCACTAAACCAAG	514
QY	474	GCTTCATTTTC	484
Db	513	CTTCTTTTC	503
RESULT 39	AE004745		
LOCUS	AE004745	11884 bp	DNA linear BCT 30-AUG-2000
DEFINITION	Pseudomonas aeruginosa PA01, section 306 of 529 of the complete genome.		
ACCESSION	AE004745	AE004091	
VERSION	AE004745.1	GI:9949336	
KEYWORDS	Pseudomonas aeruginosa.		
SOURCE	Pseudomonas aeruginosa		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas		
REFERENCE	1 (bases 1 to 11884)		
AUTHORS	Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, R.L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., and Paulsen, I.F.		
TITLE	Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen		
JOURNAL	Nature 406 (6799), 959-964 (2000)		
MEDLINE	20437337		
PUBMED	10984043		
REFERENCE	2 (bases 1 to 11884)		
AUTHORS	Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, R.L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K., Wu, Z., Paulsen, I.F., Reizer, J., Salier, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA		
FEATURES	Location/Qualifiers		
source	1. .11884		
	/organism="Pseudomonas aeruginosa"		
	/strain="PA01"		
	/db_xref="taxon:287"		
gene	102. 419		
CDS	/gene="PA3216"		
	102. .419		
	/gene="PA3216"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AG06604.1"		
	/db_xref="GI:9949337"		
	/translation="MSSPDERFHSFAEFPYPIQHSNAVCRRLLHYVGLVLAFLF		
	HAVASQQWLLALPLVGYGFAWVGHFVEKNRPATFKYPLNSFGMDWMLKDAFTGR		
	IRF"		
gene	636. 2027		
CDS	/gene="PA3217"		
	636. .2027		
	/gene="PA3217"		
	/codon_start=1		
	/transl_table=11		
	/product="probable adenylate cyclase"		
	/protein_id="AG06605.1"		
	/db_xref="GI:9949338"		
	/translation="MKPTLPDRSPTSRVSAMREYYSRLVAYTACGASIAAGPYTQYF		
	SYGLMWPYALLYPHLAYHIGFRQHPDKVTRALLAVDAVHCGLMGLAGFVVP		
	SLMFLVLSFTALVIGGLRLGLMALLVSASSALLVAVLVPPLGTPPVEVAVSTLF		
	CGLYICITAFEGHQQLRAQVQREIAREQEKARLNALAKYLSQVQWEMTFSGKS		
	VRLETORKLITVFSIDIRGTSELEAEALTDLLNNYLNEMSKIALKYIGTIDKFV		
	GDCVMVFDPSTQGAKKADAAVAVSGMIARRKIMKVLROQWRAQGIKPKLEIRMGINT		
	GCTVGNFGADTRMDYIIIGREVNLASRESASEAGEIILISHETYSLIKIDVIMCRDKG		
	QIAVKGFSRPQVIQYVWDSRRDLGAAPSYVEHELPGFSMYLDTNNIQTNDKERVIOAL		
	QQAERLRDKVIL"		
gene	2569. .3042		
CDS	/gene="PA3218"		
	2569. .3042		
	/gene="PA3218"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AG06606.1"		
	/db_xref="GI:9949339"		
	/translation="MEQFGIDALAPQVLEAPGAELVAQAARGAHQHRPARPMEGAQE		
	GIAPQQRPTAVTVQFGKARMDRGEVAAALQAPAPRGPAERLPAGQVHGIRGEGID		
	QPAEAPEVPGQDFRIAGTQAAAEVGGRHPLYVAQQFQLALQOQAGDDVDLR"		
gene	complement(3069. .3881)		
CDS	/gene="PA3219"		
	complement(3069. .3881)		
	/gene="PA3219"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AG06607.1"		
	/db_xref="GI:9949340"		
	/translation="MSTAQRFTSNRKQVRVTLWISDVHLGTRDCQAEHLAAFLKRYHA		
	DRYLYVDIIDGKWKRGYVPOAHTNVRRLTLTKRGTEVIYVGNHDEFURRYS		
	LILGNIRLVDEAVHSTADGRKLVTHGDOFVITRYHRWLAFIAGDSAYEPTLLNRL		
	NHWRBGGYGSLSAYLKHVKYAVNFTSDFEETAHSCHEKGLDGVGCGHHAEL		
	RRIEGVDYLNCGDWVESCTALIEHWGSIQLYLAEEQARLAAQAAAVEPAA"		
gene	4028. .4792		
CDS	/gene="PA3220"		
	4028. .4792		
	/gene="PA3220"		
	/codon_start=1		
	/transl_table=11		
	/product="probable transcriptional regulator"		
	/protein_id="AG06608.1"		
	/db_xref="GI:9949341"		
	/translation="MGSILSRHSHYSHDQIVHSHDHAQLVLGLSGCLDFEYEGRSVL		
	ROTFVAVPAQARHACSSPGSRCLVLDPCDGLSLGLGHEAGARLRPRPALSIIT		
	PAGQLVNLASGPIINDPVTAHQAGLLASLASERQPPQRPRLPLASIDAYLDHRA		
	APHQVADLARLAGLSVARFHSLEFLAETGOTPMDYANSRRLHQAEHLLESSLAVGEI		
	ARVGVASQSAFTAAVLRFGVTPPQLRRESRDKTG"		
gene	4857. .5192		
CDS	/gene="CsaA"		
	/note="PA3221"		
	4857. .5192		
	/gene="CsaA"		
	/codon_start=1		
	/transl_table=11		
	/product="CsaA protein"		
	/protein_id="AG06609.1"		
	/db_xref="GI:9949342"		
	/translation="MDIEMQDFERVELRVGTITLSAAPNPKARKKAPVYLVEDLGLGL		
	RTSSAQLTAYHQAELVGRQVLCVCFAPQIAGRVSEVLVTGVTYDSEKVLVLAGDFK		
	PLPNGARLA"		
gene	5202. .6101		
CDS	/gene="PA3222"		
	5202. .6101		
	/gene="PA3222"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AG06610.1"		
	/db_xref="GI:9949343"		
	/translation="MNRSSALAAHGGALLFGTGVFGKLASPSGVIVFGRALFAVL		
	AGLGFARLQAQWRALGPAGWLCGLLGHVWSFFTAIVTAGVATLAFGSPF		
	AFVILLEGILIFRERIHLECLLVLSGLVLTVPNFDLASGATGLLWAILGSLFFA		
	LISLANSQGVGPQAAWQNLVVAICLLFPVSELPAIPALDMLWALLGCVFTG		
	LAHSLVASLAVVKAETAAVLFALFPYVGTAFANILLFHPETPGPRMFLGGLIILAIVL		
	SARLGARGRTAVGDERGDPERA"		
gene	complement(6206. .6847)		


```

/notes="Function Code:14.01 - Unknown, Conserved protein;
similar to, gp:GI:gl666509 LN:LIU61226, p(-)-8.6E-16,
pid=23%"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AAB84678.1"
/db_xref="GI:2621215"
/translation="MPETIITPNLNGVSFLVYFRSLRQTYEDFRVVFIDNGSSDG
SDFIKNSSEYLGDRRIILEHENHGFASNNQGIADLDDPQCRVYVCLNNDTEVL
PDFIERLLAGACACLRVSGVQARIMWQNRKLIDSAGLEYSLNGLGNRGAYPEVNF
NERARILCAGACACLRVSGVQARIMWQNRKLIDSAGLEYSLNGLGNRGAYPEVNF
PDATVHYKGTNPETSDFTVYHNWNTWTETKNMPADLLMLPLFVLTETLIQVLL
NIRKGLIILRAXDAYGLGRVIRRRALKRPVSEIEKYLTRWSVKVPSPQED"
complement(2945..4093)
/gene="MTH173"
CDS
complement(2945..4093)
/notes="MTH173"
/notes="Function Code:8.02 - Metabolism of Complex Lipids,
Inositol phosphate metabolism ; similar to, pir:LN:F64500
AC:F64500, p(-)-9.6E-29, pid=27%"
/codon_start=1
/transl_table=11
/product="LPS biosynthesis Rfbu related protein"
/protein_id="AAB84679.1"
/db_xref="GI:2621216"
/translation="MGPMRLIVDSFFVPHYNGGERRVFEIARRLVERGHVVVDVISM
GHVGVEYEVSGRVHILGPRIKPLRGLDFIRFMAAFRWVWTHDYDIDAQTY
APLLPAFLASRIQTPWATIDHVSAAHGQWQSSKATILERVLMRLPYDQVITVS
RSTASALTLEHGRNPDGHIIPNGVDPDLDSVTPATGNVILFVGLAPKHVDHLE
VSKLVDPPDLLEILIGDVERARKAMWDEGIRDSVTFHNLSPYSEVSRKGR
VIVLPSTRGFCVNLBACAGCPAVAYSGGVVEVDGGENGLFVPCDKEALHDKI
KLISDLELRDRMSQGRKKVEEFIDWRVVDVERTYFIIARKNTP"
4263..6620
/gene="MTH174"
CDS
4263..6620
/notes="MTH174"
/notes="Function Code:12.12 - Cell Processes, Broad
regulatory functions ; similar to, gp:GI:gl652472
LN:D90905, p(-)-1.1E-36, pid=17%"
/codon_start=1
/transl_table=11
/product="sensory transduction histidine kinase"
/protein_id="AAB84680.1"
/db_xref="GI:2621217"
/translation="MNPMAIISTAFVVALVGVYIFVKNPESFRNALLLSALIAY
LSITDFGLYISENMADFWVRGFLWPVVPVILHMLLFTGRRPTPLAVIYAPA
VTAALLETDLITAGTMTSGWTFSAENSLPYOISNLWGFIIASVALIILVHWR
SRGSDRTQAHVLLTGLILPVLAVLTDLVLPAGNAVPSLTDPSAALGVSIITAYGM
RRLPLSLNAMDIVRNNTFLIITDGLNIHYNPATERLITGRRAEFGSKVTE
VLFDELSTGIMESLRKDEWIPVIVSAGHVHSGGHEGLGLLFTGSDMRPILMEE
KLRRREDRLVLETHMADGLGEFDRFNFLYSPSKKITGYDHEMLTGMALSDFMDY
LHPDDTEYVKLILTRGEGRATFRRPDGNYRWVEYVDRPIRTDGDIEGVVGLRD
IHEKLEALIVSKKRELFNMSGIIAVERCKGFRVSKMNPAAERMEGVSCAA
ICQMDIQEALPWNPSPIRRAQLEIBETGEALHCPVECEGWVHLQLSGEVVI
ICTDITLRQYEDELASLREKALLREVHRVKNQFQVSSINLQDDAEDAPLR
DQSRQSMALWHELYSEDLTSDMGRIEKLTSIIVNSHNGIEIEVAVGDITL
PLETALPLGLINELNYSTFKHFTSGMYSIELEHGGFEFTLIIDNGLVGLPPDFII
EDSDSLGLRLVAGVQIDGTLEVSGEDGTRFRLTFRGVVYPRRV"
complement(6003..6074)
/repeat_region
/notes="72 bp direct repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
83% ID to interval 400218"
/repeat_region
complement(6010..6094)
/notes="85 bp direct repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
83% ID to interval 228380"
/repeat_region
6011..6089
/notes="79 bp inverted repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
86% ID to interval 4087"
/repeat_region
complement(6011..6089)
/notes="79 bp inverted repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
86% ID to interval 4086"
complement(6011..6082)
/notes="72 bp direct repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
86% ID to interval 898474"
complement(6015..6082)
/notes="68 bp direct repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
87% ID to interval 384764"
6020..6101
/notes="82 bp inverted repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
79% ID to interval 3058"
complement(6020..6101)
/notes="82 bp inverted repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
79% ID to interval 3057"
complement(6020..6082)
/notes="63 bp direct repeat includes part of MTH174
(sensory transduction histidine kinase - paralog.fam. 1);
86% ID to interval 554696"
6749..7945
/gene="MTH175"
CDS
6749..7945
/notes="MTH175"
/notes="Function Code:14.01 - Unknown, Conserved protein;
similar to, pir:LN:F64383 AC:F64383, p(-)-1.4E-41, pid=27%"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AAB84681.1"
/db_xref="GI:2621218"
/translation="MALTPRGDVLRTVMYLETESRDSVDRRLWRSISDKRTIRAD
GSYTLKSEAGEAMHTRVGTALTEAEKVFRTVSGIKDDLVLDCSLGNTAALEAF
TDASSVDMVNEISACTMAREYLPSPNPAHKVYKAYEDFLLEGLSLRTSPPPAP
GIELRVHGDAVAVPLLEREYIDALFAPSPRVAPELYTVEFISLAGVLKRGVL
ATYSAAPRAALIEAGFHVGGPFGKSGTTLASMDPRMVKPLDWRDERMIALSD
AGIPYRDELSADALEIMGRVRRMERSARGVTRISSAVKTPLYLGSEOVGRGRRVR
RNLGRIGIDDPGSGPEALYIICPQDECICGCGEDRPPASSRDRVISMRRRLMDLVNLRH
LADKAG"
8114..10111
/gene="MTH176"
CDS
8114..10111
/notes="MTH176"
/notes="Function Code:10.07 - Metabolism of Macromolecules,
Aminoacyl tRNA synthetases and tRNA modification ; similar
to, pir:LN:D64354 AC:D64354, p(-)-1.4E-121, pid=37%"
/codon_start=1
/transl_table=11
/product="trna-guanine transglycosylase"
/protein_id="AAB84682.1"
/db_xref="GI:2621219"
/translation="MISLSRGNTLLMTLWEDPMFEIKSKDGLGRTGILKTEHGTVRT
PALMPVHPGKQIDVKGPGAEIITNAVITYNPNRERALSDDGVRHRLIEFDGPIMT
DSGSFOLSGDIEVENPFIIRFQEDIGTDSLDITPPPGVSHRRLAEVETPLER
ARESIYRERMLNAVYQSTHPDLRYCASRLAEPLVHLPAGVAVPLMEVYREL
VDVALSVSELPPSRPHLMGAGHPMLFALAVSMGCDLFDLSAAYILVAEDDRLLSTEG
TYKLENLQEMPCSCSVCTDTPSELMDREERNLIAENLHVFAERKVRQAIDH
GNLMELVEERCAHRAHLEGIRYRMSEYLDLIEKFEPRSKRSAFFYTGPSLGRVVEHR
HLKRVKHELGERLALVAPRYPYSSLPARIGFSSLRPQSGGPRWVWVLDPLFGIIP
LELDQVYPLAQSDAPGIMDLGDEEFLGLVRDLMGDGDAIVDDALCSELGTELPKYGM
EVETTVDLDRVMADYQFGMGAGELLFTDDVIERSTNGTIRIHYAGDELICTMR
ASDGLLVLAGAVARLHKGTDPYPAWVAVNEESEPARKGKSVFAKFIIDCDNNIRAN

```

Query Match 3.9%; Score 46; DB 1; Length 11142;

Best Local Similarity 56.7%; Pred. No. 0.0095;

Matches 85; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 366 AACAGCGTCCCTGCACATCTCTTCGNATGCATCAGGATCCGCCCGCTCTCAGTGGCG 425

Db 3129 ACCAGAAACCGGTTTTCACCGTCATCTATGACCTCAACACTCCACCGGCTGTAGGCC 3188

Qy 426 ATACGGGACGCGGAGACTGACCTTCAGCGACGACATACCAACACCTTCATTTC 485
 Db 3189 ACAGGACGACACACGACGACCTTCAGCGACGACATACCAACACCTTCCTG 3248
 Qy 486 GAAGGATGACACACGACGACCTTCAGCGACGACATACCAACACCTTCCTG 515
 Db 3249 GATGGAAGGACACGACGACCTTCAGCGACGACATACCAACACCTTCCTG 3278

RESULT 41
 AF498404/c 13401 bp DNA linear BCT 13-JUN-2002
 DEFINITION Pseudomonas aeruginosa serotype 013 putative O-antigen biosynthesis
 gene cluster, partial sequence.
 ACCESSION AF498404 AC104723
 VERSION AF498404.1 GI:20559835
 KEYWORDS
 SOURCE
 ORGANISM Pseudomonas aeruginosa.
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 Pseudomonas.
 REFERENCE 1 (bases 1 to 13401)
 AUTHORS Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
 Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
 TITLE Genetic variation at the O-antigen biosynthetic locus in
 Pseudomonas aeruginosa
 JOURNAL J. Bacteriol. 184 (13), 3614-3622 (2002)
 MEDLINE 22053227
 PUBMED 12057956
 REFERENCE 2 (bases 1 to 13401)
 AUTHORS Raymond,C.K., Sims,E.H., Kas,A., Spencer,D.H., Kutayavin,T.V.,
 Ivey,R.G., Zhou,Y., Kaul,R., Clendinning,J.B. and Olson,M.V.
 TITLE Direct Submission
 JOURNAL Submitted (04-APR-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98105-2145, USA
 COMMENT On May 13, 2002 this sequence version replaced gi:17975252.
 FEATURES
 source
 1..13401
 /organism="Pseudomonas aeruginosa"
 /serotype="013"
 /db_xref="taxon:287"
 <1..104
 /note="ORF_1: C-terminal coding region of rpsA"
 /codon_start=3
 /transl_table=11
 /product="RpsA"
 /protein_id="AA027608.1"
 /db_xref="GI:20559836"
 /translation="EKDANKELKQEVESAGPTTIGDLIRAQWENQG"
 241..525
 /note="ORF_2: hmd/ihfB"
 /codon_start=1
 /transl_table=11
 /product="Hmd"
 /protein_id="AA027609.1"
 /db_xref="GI:20559837"
 /translation="MTKSELIERIVTHOGLSAKDVLAIKTMLEQMSQALATGDRIE
 IRGFGFSLHYRAPRVGRNPKTGESVRLDGKVPFKPGKELDRVNEPE"
 597..815
 /note="ORF_3: hypothetical orfA"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027610.1"
 /db_xref="GI:20559838"
 /translation="WVALTVIFVLENQGVNLSFLGWRTPGPLSLFVSLAFVLGGL
 GGVLSGMWLRLLIGSVSKGHLVPRKAM"
 1167..2216
 /note="ORF_4: wzz; similar to chain length determinant
 protein"
 /codon_start=1
 /transl_table=11
 /product="Wzz"
 /protein_id="AA027611.1"

/db_xref="GI:20559839"
 /translation="MTEETFFQVAKNEEDFLGLRLRLMAEKWLIACTITLITSLVG
 TYALSPVYEAIRAVPPSPSAGNLAENGLATTSVNDVYSVFLRNLLAEETR
 NRRFVYVPSLSERSGSKDRYSNLSEMEIQRPQSDRERYTVIFEKHDGQAA
 DWGRYIKLAAEGSKEMQDAVREFEVKSNIOQDILLDRADAKAREINFLKKA
 YQIAESLKENPPLIGGMDQQLSSIMEGALMYNKGTKALRAELAEERTSDPPFP
 ALRNQEQAMYSTLKLDVNVKVAFRQDQVAEVDPPVRPKKLLLVGLGSLVWVG
 MIAFVAVRRSRK"
 2303..3568
 /note="ORF_5: similar to UDP-glucose/GDP-mannose
 dehydrogenase"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027612.1"
 /db_xref="GI:20559840"
 /translation="MDIKAAVGLVGLPLAVEFGKRSVTGFDINHRSIRDLQKY
 DSTLEVEKEELSAVHLKFTSTLSELOECNEFFIVTPTIDEHKQDPLPLVKAESI
 AKVLKNDIVYESTVPGATEEVCVPLERESGLRNSDFVGYSPERINPGKEHR
 VSTIKKVTSGTPEAEIVDSLYREIITAGTHKAESIKVAEAKVINTORDNLIALI
 NELAIIFNKLEIDTESVLAAGTKNFIPLPFGVGGHCGVDYPIYTHKAQSTGYHP
 EIIAGRLNDMGAYVVSQVKAMLRKRIHVDGARVLVGLTEKNCPLDNRNKVD
 IVRLAEYNIADVDPVMSVEEAQHEVGLTPISAPVEGNYDAVVLAVAHNEFKELGA
 DKIRAFGKLESVIYDLKFLDKSDSLR"
 3578..4603
 /note="ORF_6: similar to NAD dependent
 epimerase/dehydratase family"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027613.1"
 /db_xref="GI:20559841"
 /translation="MNAKALSRLETTSSWLVTGCVAGFTGSLNLELLKFNQKVVG
 LDDFTGHMNLDEVRGLVEERWRNFKIQGDIRDIDCRRAYEGVDYVHLHQALGS
 VPSIADPTTNSVNIISGFLNLAKEAGVKSILTYAASSYTGHDGPKLKEVSIK
 PLSYAVTKYNELIADVFSRCYGNLIGRYFNVGKRODPNGAYAAVLPKWSAM
 KPEVYINGDGETSRDCCFINNVIKANLAAFCETEAQKNOVNVAYGGRTLGQDP
 SLKVELSBNGINYGQPIYREFRAGDVYHSAQEIKKQAQVLYGNYPEYDIQAGIAEMP
 WYISFVK"
 4665..5954
 /note="ORF_7: similar to Polysaccharide biosynthesis
 protein; potential multiple membrane spanning domains."
 /codon_start=1
 /transl_table=11
 /protein_id="AA027614.1"
 /db_xref="GI:20559842"
 /translation="MNPKNIEWNLGLPLIIAALVPHLIATVVEREGFLALSNG
 LIGYALDILGIGRAVTKLSSIRGSEEMNSYTIMTAIRTAIVGGLGFLIIIFG
 LIGGEFFSHKDVSSLEFVLLALFALPQAI SATYRGVNEAYLNFKGNLRIFL
 GATNFGPFLSVYAKOLHYLVTLVLSRALAFFERLHAGVLEKTELVERSCYDRR
 QAVELFRGGHWTVSSINPELVSDRFFIGVLLSAAAVSYVPIEYITQSMILVGA
 VSTVAFPSISNLRTSFALAEFCFNKLLRVLLIMGGMLCLAFLLPILKLVWGDYI
 GDSISVGRILCLGVFFNALGMFYSLHANAKVKEITAILHSIELPILILILIPR
 HGIYGAVAWSLRLTVDITVLAVLSYFGWRVKEANC"
 6001..7176
 /note="ORF_8: similar to Sodium/hydrogen exchanger family;
 potential multiple membrane spanning domains."
 /codon_start=1
 /transl_table=11
 /protein_id="AA027615.1"
 /db_xref="GI:20559843"
 /translation="MVETALIGIKETILLIALLFTLISSLARADNLVYLLAGLFS
 VGVLWMSYIGELRGNAGALQVTVVTVPLVLLAVSYRQSEALGKLPFYAALLIA
 YLIDFFFGNYLFSYNPYVESMAAYGDAVVDNAGRYIKFTLPNVSSLIFLPYVVS
 YLIDKRFSLGLFLALLVLLVSSAGRALLVAFTAPFLAYFLTMGADRQFWYVS
 VLLFVFMGLFLIFYNPEYVENLFLSISNDSNERRRYQFALMKGVYQOQP
 FGMAAGDAIVRSEEMWYELFVSLVFOYGLIGLFYALGVLELLVFLVLRVKI
 GRSPFCYLSGMLAFMATATNPYLAQFDMWVIFIPVALLNYSVSRGKDGNSY"
 7157..8092
 /note="ORF_9: similar to Glycosyl transferase"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027616.1"
 /db_xref="GI:20559844"
 /translation="NESFIIDVVVWNNAGLQLRACVDSIFKSSGLIRKIVVIDNGS

Matches 150; Conservative 0; Mismatches 152; Indels 3; Gaps 1;
Qy 329 CAGTCCCGGTTCATCCACACATCCATCCCGAGGTAAACAGCGTCCCTGCACATCTTT 388
Dy 225825 CATGCCCGGGCAGCGCGCGGCATCCCTACCGAACAAAGCGCGCAACCGCCCTC 225766
Qy 389 CTGAATGA--CATCAGGGATCCGCCCGCTCTCACTGGCATACGGGCACGCGGAGAC 445
Dy 225765 CGCAAGATTTCGCCCGAGCCCGAGGGCGAGTGGTGAACACGCGGAGCGGAAGGA 225706
Qy 446 TGACGCTTCAGCCAGTACCATACCAACACGCTTCATTTCCGAAGGATGACCAACCACT 505
Dy 225705 CAGGCGCTCAGCATACGTTGCCAAGCCTTCATATCGGAGGACAGGACGACATC 225646
Qy 506 GCAATCCGCTAGACCGGTAACTGGAAGGACCGTCCATTAACATCACTCCGCT 565
Dy 225645 CGCGTGGCTAGAGGCTGAGATCGGGGTGGAAGCCAGCAATACCGGCTGC 225586
Qy 566 CATTCACAGTGTCTCTGCTGACGACGACGCTGCTTCATTTTCACGCGCGGCGCC 625
Dy 225585 AATCCGAGCTCGGCGCCAAAGCGCGCAAGGACGCGCTCGTTGACCCCTCCCGAGGAG 225526
Qy 626 CACCA 630
Dy 225525 CATCA 225521

RESULT 43
AE0000469 10940 bp DNA linear BCT 01-DEC-2000
LOCUS Escherichia coli K12 MG1655 section 359 of 400 of the complete
DEFINITION genome.
ACCESSION AE0000469 U00096
VERSION AE0000469.1 GI:1790385
KEYWORDS
SOURCE Escherichia coli K12.
ORGANISM Escherichia coli K12.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 10940)
Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
Mau, B. and Shao, Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (5331), 1453-1474 (1997)
97426617
PUBMED 9278503
REFERENCE 2 (bases 1 to 10940)
Blattner, F.R.
Direct Submission
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
3 (bases 1 to 10940)
Blattner, F.R.
Direct Submission
Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
4 (bases 1 to 10940)
Plunkett, G. III.
Direct Submission
Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, Director).
Supported by NIH grants HG00301 and HG01428 (from the Human Genome
Project and NCHGR). The entire sequence was independently
determined from E. coli K12 strain MG1655. Predicted open reading
frames were determined using GeneMark software, kindly supplied by

Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
30332 [e-mail: mark@ambr.gatech.edu]. Open reading frames that
have been correlated with genetic loci are being annotated with CG
Site Nos., unique ID nos. for the genes in the E. coli Genetic
Stock Center (CGSC) database at Yale University, kindly supplied by
Mary Berlyn. A public version of the database is accessible
(http://cgsc.biology.yale.edu). Annotation of the genome is an
ongoing task whose goal is to make the genome sequence more useful
by correlating it with other data. Comments to the authors are
appreciated. Updated information will be available at the E. coli
Genome Project's World Wide Web site
(http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
its annotations are periodically updated; this is version M54. No
sequence changes. Annotation updates: updated gene identifications
and products; all new functional assignments courtesy of Monica
Riley; added promoters, protein binding sites, and repeated
sequences described in reference 1. The unique numeric identifiers
beginning with a lowercase 'b' assigned to each gene (protein- or
RNA-encoding) are now designated as gene synonyms instead of
labels. This should allow them to be searched for in Entrez as gene
names.

FEATURES
source
1. 10940
/organism="Escherichia coli K12"
/strain="K12"
/sub_strain="MG1655"
/db_xref="taxon:83333"
95. .114
/note="central position to predicted promoter: -145"
/bound_molety="CRP predicted site"
215. 244
/note="factor Sigma70; predicted +1 start at 4140062"
298. .1377
/gene="frwC"
/note="b3949"
298. .1377
/gene="frwC"
/function="transport; transport of small molecules:
carbohydrates, organic acids, alcohols"
/note="o359; 100 pct identical amino acid sequence and
equal length to PTWC_ECOLI SW: P32672"
/codon_start=1
/transl_table=11
/product="PTS system, fructose-like enzyme II component"
/protein_id="AAC76931.1"
/db_xref="GI:1790386"
/translation="MNELVQILKNTQHLMTGVSHMIPFVVSGLLAVSVMLYKGA
VPDAVADPNLKLFDIGVAGLTLMVPLAAVIGVSIASALAPCAIGAWGNFSGAG
FFGALTAGIIGGIIVVHYLKKIPVHKLVSRVPIPIVGTLTITAGIMMGICGPVGA
LTNSLTQWLOGMQGGSIVMLAVINGMLAFDMGPGPNKVAFAMLICVAGVITVVAI
AAVGCICPPLGMGLATLIGRNFSAEERTGKAALVMGCVGTGEGATPFAAADPLRVI
PSIMVSGVCGAVTAALVGAQCYAGWGGLIPLVPEVGEKLGITAAVAVGAVTAVCVNVL
KSLARKNGSTDEKEDDLDFEIN"
1392. .1712
/gene="frwB"
/note="b3950"
1392. .1712
/gene="frwB"
/function="enzyme; Degradation of small molecules: Carbon
compounds"
/note="o106; 100 pct identical amino acid sequence and
equal length to PTWB_ECOLI SW: P32673"
/codon_start=1
/transl_table=11
/product="PTS system fructose-like IIB component 1"
/protein_id="AAC76932.1"
/db_xref="GI:1790387"
/translation="MTKIIATVACPQSGVAHTYMAAEALSAKAKGEVETQGSIG
LENELIAEDVASADWMLTKDIGIKFEERFAGKTIVRVNITSDAVKRAATMSKEAHL
AQA"
1598. .1625
/gene="frwB"
/note="factor Sigma70; predicted +1 start at 4141443"

promoter
1684. .1715
/note="factor Sigma70; predicted +1 start at 4141533"
gene
1763. .4060
/gene="pfid"
/note="b3951"
1763. .4060
/gene="pfid"
CDS
/EC number="2.3.1.54"
/function="enzyme; Energy metabolism, carbon: Anaerobic
respiration"
/note="o765; 100 pct identical amino acid sequence and
equal length to PFLD_ECOLI SW: P32674"
/codon_start=1
/transl_table=11
/product="formate acetyltransferase 2"
/protein_id="AAC76933.1"
/db_xref="GI:1790388"
/translation="MTNRSRLKTAFTANTREISLERALLYTASHRQTEGEPIVLRRA
KATAYILHEVEISIRDEELIAGNRVTKPRAGIMSPENDPYLLKELDQPTTPQDRFA
ISEDKRIYREELFPYWEKRMKDFNGQMTDEVKAATNTQIFSIQTDKGGHIIID
YPLLNHGIELVAOQHCQOQOPENHFYQALLLEASOKHILRYAEALAEYMAANCT
DAORRELLTIAELSRNQHOKPQTFWQACOLPYMNIILQYESNASSISLGRFOQYM
LPYQISLQGEADAELFELLESILWVKNCNDVLLRSTSSARFAGPTGYTALLGLT
ENGRSAVNSLFCLDAYQSVLPQPNLGVRTNALIDTFLMKTAETIRFGTGPQIF
NDEVVPAFLNRGVSLDARDYQVGCVELSPGTYGLHDIAFMFLKLVMEICLHEN
EGNAALTYEGLLEQIRAKISHYITLWVEGSDICIDIGHRDWPVPLLSFISDCLKGR
DITDGGARYNSFGVQIGIANLSDSLHAKGMVEEQRLSFEDELLSVLKANFATPEGE
KVARLNRKPEKGNIDVDNISALLRHCKVEKYNPRGGYFTPGSYVSAHVHP
LGSVGTAPPRAGAGLADGGLSPMLQDQAGPTAVLKSVDKLDNTLLSNTGLLVK
FTPALLEGAGLRKLDLFLRAFTQLKQHQIFQNVNADTLREAAQRPQDIAGLVLRVA
GYSAFFVELSKETODDIIRTAHOL"
3292. .3329
/note="REP (repetitive extragenic palindromic) element;
contains 1 REP sequence"
4026. .4904
/gene="pfic"
/note="b3952"
4026. .4904
/gene="pfic"
/EC number="1.97.1.4"
/function="putative enzyme; Energy metabolism, carbon:
Anaerobic respiration"
/note="o292; 100 pct identical to PFLC_ECOLI SW: P32675"
/codon_start=1
/transl_table=11
/product="probable pyruvate formate lyase activating
enzyme 2"
/protein_id="AAC76934.1"
/db_xref="GI:1790389"
/translation="MTSSAGORISCNVETRRDDVARIFNIQRYSLNDGEGITVVVFF
KGCPLCPWCANPESISGKTQTVRRRAKCLHCAKLRDADPCPSGAFERIGRDISLDA
LERVMKDDIFFFTSGGVTLGSGEVLMOAEFATRLQRLRWGSCATETAGDAPAS
KLPLAKLCEVLFDLKDMDATQARDVKNMLPRVLENRLRLVSEGVNVPRLPLIPG
FTLSRENQALDVLPLNIRQHLPLFPHQYGEKPYLLGKTSMEKEVPAPSSADVAT
AREMAERAGLQVTVGG"
4730. .4760
/gene="pfic"
/note="factor Sigma70; predicted +1 start at 4144578"
4864. .4894
/gene="pfic"
/note="factor Sigma70; predicted +1 start at 4144712"
4906. .5247
/gene="frwd"
/note="b3953"
4906. .5247
/gene="frwd"
/function="enzyme; Degradation of small molecules: Carbon
compounds"
/note="o113; 100 pct identical amino acid sequence and
equal length to PTWX_ECOLI SW: P32676"
/codon_start=1
/transl_table=11
/product="PTS system fructose-like IIB component 2"

/protein_id="AAC76935.1"
/db_xref="GI:1790390"
/translation="MAYLVATACVSGVHAHTYMAAERLEKLCLEKWCVSTETGALG
TENRLADEDIRADVALITDIELAGAEERFECRYVQCSIYAFLEPQVMSAVRKVL
SAPQOHLILE"
complement(5234. .6085)
/gene="yij0"
/note="b3954"
complement(5234. .6085)
/gene="yij0"
Query Match 3.8%; Score 45.4; DB 1; Length 10940;
Best Local Similarity 63.1%; Pred. No. 0.015;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 11 TCGAAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
DB 7949 TAGAGAAAAACCAAAAGTAGAGACGTGCCAACCCAGGCGCTTCAGCTAAAAAGAGGTT 8008
QY 71 TGTCTGTCGGGTATTTAAATGCATTCACCGTCCCGCTTAAACAATGT 121
DB 8009 TAGCCTGGACTTCTCTGGAATCATAGTAAGAACCTGCTTTGAAAAAATAT 8059
RESULT 44
LMFLCHR32_19
WPCOMMENT
Sequence split into 28 fragments LOCUS LMFLCHR32 Accession AL499622
Fragment Name Begin End
LMFLCHR32_00 1 110000
LMFLCHR32_01 100001 210000
LMFLCHR32_02 200001 310000
LMFLCHR32_03 300001 410000
LMFLCHR32_04 400001 510000
LMFLCHR32_05 500001 610000
LMFLCHR32_06 600001 710000
LMFLCHR32_07 700001 810000
LMFLCHR32_08 800001 910000
LMFLCHR32_09 900001 1010000
LMFLCHR32_10 1000001 1110000
LMFLCHR32_11 1100001 1210000
LMFLCHR32_12 1200001 1310000
LMFLCHR32_13 1300001 1410000
LMFLCHR32_14 1400001 1510000
LMFLCHR32_15 1500001 1610000
LMFLCHR32_16 1600001 1710000
LMFLCHR32_17 1700001 1810000
LMFLCHR32_18 1800001 1910000
LMFLCHR32_19 1900001 2010000
LMFLCHR32_20 2000001 2110000
LMFLCHR32_21 2100001 2210000
LMFLCHR32_22 2200001 2310000
LMFLCHR32_23 2300001 2410000
LMFLCHR32_24 2400001 2510000
LMFLCHR32_25 2500001 2610000
LMFLCHR32_26 2600001 2710000
LMFLCHR32_27 2700001 272709
Continuation (20 of 28) of LMFLCHR32 from base 1900001 (AL499622 Leishmania major chr
Query Match 3.8%; Score 45.4; DB 2; Length 110000;
Best Local Similarity 51.2%; Pred. No. 0.025;
Matches 106; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 337 CGGTCCATCAGACACATCCCCACGGGTAAACAGGTCCTCTCACATTTCTTCTGAATGA 396
DB 94357 CGGCCAGTCGGCGCGGATGCTGGGGCGCACCAAGTAGCGGTGCGCGCTTCAGCACGG 94416
QY 397 CATCAGGGATCCGGCCGCTCTCCTGCGGATACAGGGACGCCGAGACTGACGCTTAG 456
DB 94417 TGTGCGGATGCGCGGACCGCGGCGGCGGCGGATCAGCGGTACGCGGCGCATCGCTCGA 94476
QY 457 CCAGTACCATTACCAACAGCTTATTTCCGAGGCGATGACCAACACTGACACTGGAATCCGCT 516

Db 94477 CCGCGTGTATGCCGAACGCTGTACCATTGGTGTTGCAGCAACACATCGGCCGCGCTGT 94536

QY 517 AGACCGGTAACGCTGGAAAGGCAC 543
||||| | | | | | | |
Db 94537 AGAACGCGCAGCTGGTCACGCCGC 94563

RESULT 45
ECOW89 176195 bp DNA linear BCT 17-DEC-1993
LOCUS E. coli chromosomal region from 89.2 to 92.8 minutes.
DEFINITION U00006
ACCESSION U00006.1 GI:409785
VERSION
KEYWORDS
SOURCE Escherichia coli (sub_strain MG1655, strain K-12) (library: lambda)
DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 176195)
AUTHORS Blattner,F.R., Burland,V., Plunkett,G. III, Sofia,H.J. and Daniels,D.I.
TITLE Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes
JOURNAL Nucleic Acids Res. 21 (23), 5408-5417 (1993)
MEDLINE 94089392
PUBMED 8265357
REFERENCE 2 (bases 1 to 176195)
AUTHORS Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1993) 608-263-7459
COMMENT On Oct 29, 1993 this sequence version replaced gi:396288.
This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award HG00301 from the NIH Human Genome Project. The entire sequence was independently determined from E. coli MG1655; overlaps and conflicts with other sequence determinations are annotated. The start of this entry overlaps the end of the entry ECOW87 (L19201) by 93 bp. NOTE: An update was submitted on 25-OCT-1993, reflecting a correction to the heme sequence.

FEATURES
Location/Qualifiers
1..176195
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"
/note="this sequence comprises the following lambda clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203, EC18-105, EC18-110, EC17-200, EC18-33, EC30-262, EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27; M13mp19 or Janus vectors were used for subcloning"
misc_feature <1..1737
/note="corresponds to M21516; ECORATGA(1225..2805)"
gene 1..1257
/gene="katG"
CDS <1..1257
/gene="katG"
/EC_number="1.11.1.6"
/standard_name="catalase HPI; catalase-peroxide; catalase"
/note="CG Site No. 14983"
/codon_start=1
/transl_table=11
/product="catalase hydroperoxidase I"
/protein_id="AAC43048.1"
/db_xref="GI:396289"
/translation="AITSGLVEVWVTQTPTQSTWNFFNLFKYEVOTRSPAGAIOPEA VDAPIIIDPDPSKKRRPTMLVDTLRLDFEPEKISRFLNDPDQAFNAFAFXWF LHWASGTFSRGDKRGKANGARALAMPORDMVAAYRALPVLEIKOESCKSLA DIIVLAGYGVGEKAASAGLSIHPPFAPGRVDAQDQTDIEMELLEPIADGFNTYRA RDVSTTESLLIDKQAQULTPAPEANTLVGMRVILGANFGSKNGVFTRDRVGVLNSF FYNLDLYEWKATDESKELPEGRDETGYVKFTASRADLVFGSNVLRVAVASS

misc_feature <1..161
/gene="katG"
/note="corresponds to lambda clone EC18-126"
misc_feature 1..93
/gene="katG"
/note="93 bp overlap with end of L19201 (ECOW87)"
misc_feature 156..15760
/note="corresponds to lambda clone EC14-54"
misc_difference 938
/gene="katG"
/note="G in M21516; C here"
terminator 1266..1305
/note="putative"
CDS 1317..2255
/note="similar to Desulfurolobus ambivalens hypoth. 28.3 kDa protein in sor 3' region"
/codon_start=1
/label_ORF_0312
/transl_table=11
/protein_id="AAC43049.1"
/db_xref="GI:409786"
/translation="MGERKLHTGSMAACKSNPLAISGLVLTLSWSYWIPIKQVT SYIGAFDTALRCIFGALVLFVILLRGRMRPTFKYITAIALLQTCGVGLAQNAL YSGGAKRYAILSYTPFWFIYAALFGERLRGQYFAITAAFGFLTQLQPVALPQS SMKSAAILSCVSGWASIAIKRYLRHPRVDLLSLTSWMLYAALVMSSVALLPVO REIDWPVFWALAYSAILATALAWSLWLFLVKNLPASIASLSTLAVPCGVLSFWLL LGENPGAVEGSCVILVIALALVSRRKEAVSKRI"

misc_difference 1429..1432
/note="TTT in M21516; TTTT here"
misc_difference 1720
/note="C in M21516; G here"
CDS complement(2282..2899)
/codon_start=1
/transl_table=11
/label_ORF_f205
/protein_id="AAC43050.1"
/db_xref="GI:396291"
/translation="MKASLALLSLTFTSHLSKSPAVPTVQIQANTNLATADGAR QOIGSTFLFPAYVOLYPCGDVPQERGCSDDVIRALRSOKVDLQKLVHEDMAKFA EYPQNKLKRPDSNIDHRVRPLETWFSRHDKTKPTSNDPSFOAGDIVSWRLONGLA HIGVVSDGFARDGTPLVIIHNIGAGAEEVDLFNMVRGHYFYVK"
promoter complement(2886..2915)
/note="promoter matrix score of 42; putative; within ORF f205, which would suggest alternate start codon"

misc_structure 1912..3053
repeat_region 3031..3064
/standard_name="REP; repetitive extragenic palindromic element"
/note="contains 1 REP sequence"
misc_structure 3090..3219
/note="predicted bend of 75 degrees"
gene complement(3174..4316)
/gene="glbA"
CDS complement(3174..4316)
/gene="glbA"
/note="Similar to Bacillus stearothermophilus glycerol dehydrogenase"
/codon_start=1
/transl_table=11
/label_ORF_f380
/product="glycerol dehydrogenase"
/protein_id="AAC43051.1"
/db_xref="GI:396292"
/translation="MPHALLSIKGAIMDIRIOSPKRYIQGADVIRNLGELYLPKLER WLWVGDKEFLVAGTEVKSFKADGLVEIAPFGESQNEIDRLGIAETAOCCGAIL GIGGCKLDTAKAHFMGVPAVIATASTDPCSALSITYTDEGFDYRLLLPNP NMVLDTVKAGAPARLIAGICDALATEARCSRSGATWAGCKTQAAALAEI CYNTLIESEKAMLAEOHVVTVPALERVIANIYLSGVFSGSGLAAAHAHVNLGTAI PDABHYHGEKAFVGTGLTQLVLENAPVEEITVAALSHVGLPITLQDLKDVEDPAK MRIVAEACAEGEITHNMPGGATPDQYVALLVADQYGRFLQEW"

misc_structure 4022..4156

```
/note="predicted bend of 75 degrees"
complement(4288..4950)
/note="similar to Bacillus subtilis hypoth. 20 kda
protein, in tsr 3' region"
/codon_start=1
/transl_table=11
/label_ORF_f220
/label_id="AAC43052.1"
/db_xref="GI:396293"
/transl_table="MELYLDNTANVAEVEERLARTFPFIAGVTTPNPSIIIAAKESISWEVLP
RLAKAIGDILFQAGMRDAQGVVEARLDAIPGIVVVKIPVTSEGAAIKLKKE
GITTGLTAVYSAAGLILALAGAKYVAPYVNRVDAQGGGIRTVQELQTLLEHAPES
MWLAASFTRQALDCLLAGCESITLPLDVAQQLMNTPAVESALIERFEHDWNAAFQTT
HL"
complement(4922..4949)
/note="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
complement(4962..7097)
/note="similar to phosphotransferase system enzyme I"
/codon_start=1
/transl_table=11
/label_ORF_f711
/label_id="AAC43053.1"
/db_xref="GI:409787"
/transl_table="MCSGSAGGILTPISLDNALGNLPAKGVDAEQSALENGLTIV
LKNTEFLDSGATSAILLRAHRSAGDTLSREHLAGVSGALSCEAIVASANHFCE
EFSRSSYLOERALDYVCFQQLQIYGEORFPAPKLTQPAICWADLETSPQFLE
LDKNHGLLLKSGTTSHTVILARSNPITLVGDIDALTQWQOTIXIDGNAGAIIV
VEPGEAVARYTQGARVODALRQORWLTQQARTADGIRIEIAANIAHSEVQAAPG
NGAGVGFLTEMLYDRTSAPGESELYNFCQALSSANGRSIIVRTMDIGGDKPVDY
LNIIPAENPFIYRAVRYEYELFTQLRSILRASAHGSIKIMIPMISSMEEILLWY
KEKLAARQQLRNHIEPFDEKIQIGIMLEVPVSMFTIDOCCEIDFSGSNDLTQYL
LAVDRNAKTRVHNSINPAFLRALDYAVQAVHROGKWIGLCELGAKGSLVPLLVGL
GIDLSKSAPSIIPAKARQOLDSRECKLIQNAACRTSLEVEHLLAQFRMTQODAP
IYTAECITLESDSKSEVLEKWTMDNLLAGRCRYPRKLEADLWAEAVSTGLGSPF
AIPHSKSHIEDSTISVARQAPVRWGDDEQFIIMTLNKHAGDQHRMIFRSRLARR
IMHEEFRLNALYAAADAIASLLQHELEL"
7099..7428
/codon_start=1
/transl_table=11
/label_ORF_0109
/label_id="AAC43054.1"
/db_xref="GI:409788"
/transl_table="MRARIICGFLVSETGGSSSLLTSASGASQWNGNSRSRLHKLK
WCACSSAPEISNQLSPASVSPISARALLALPCGCAGYATQVNMKRYTAPFQRL
PDHVR"
complement(7108..7136)
/note="promoter matrix score of 46; putative"
7200..75839
/note="corresponds to lambda clone EC21-52"
7578..7699
/note="putative"
7771..8850
/note="similar to phosphotransferase system enzyme II"
/codon_start=1
/transl_table=11
/label_ORF_0359
/label_id="AAC43055.1"
/db_xref="GI:396296"
/transl_table="NNELVOLKNTROHLMGTGVSHMIPFVYSGGILLAVSYMLYKGA
VPDADVADNKLKFLDILVAGLTVMVPLAAVIGYSAERSALAPCAVWGNSEFGAG
FGLAGTLGIGVYHLYKKIPVHKVLSRVMPFIPIPIVGLITAGIMWGLGPEVGA
LTNSLUTLQGMQCGQSVMLVAVIMGLMFLAFDGMGVNPKVAFMLICVAGQVYTVVAL
Query Match
Best Local Similarity 63.1%; Pred. No. 0.027;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 11 TCGAAAAAGGCAAAATAAAAAATGTCCTCCAGCGCGCTCCAGCTGAAGTAGGCC 70
Db 15422 TAGAGAAAAACCAAAAGTAGACAGTGCACCAACCCAGGGCTTCCAGCTAAAGAGGTT 15481
QY 71 TGTCTGTCCGGTATTATTAATGCAATGACCGTCCCGCTATTTAAACAATGT 121
```

```
Db 15482 TAGCCTGGACTTCTGTGGAATGCATAGTAAGAACCTGTCTTGAATAATAT 15532
RESULT 46
AC023609/c
LOCUS
DEFINITION
AC023609 196146 bp DNA linear HTG 29-JUN-2000
Mus musculus clone CT7-405C9, WORKING DRAFT SEQUENCE, 34 unordered
pieces.
AC023609
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 196146)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 196146)
DOE Joint Genome Institute.
Direct Submission
Submitted (16-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jun 29, 2000 this sequence version replaced gi:6980219.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1108758
Center clone name: RG-MBAC_405C9
-----
Summary Statistics
Consensus quality: 168540 bases at least Q40
Consensus quality: 179014 bases at least Q30
Consensus quality: 181723 bases at least Q20
Estimated insert size: 160000; agarose-fp estimation
Estimated insert size: 192846; sum-of-contigs estimation
Quality coverage: 6.47 in Q20 bases; agarose-fp estimation
Quality coverage: 6.47 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1073: contig of 1073 bp in length
* 1074: gap of unknown length
* 1174: contig of 1184 bp in length
* 2358: gap of unknown length
* 2458: contig of 1329 bp in length
* 3787: gap of unknown length
* 3887: contig of 1195 bp in length
* 5082: gap of unknown length
* 5182: contig of 1136 bp in length
* 6318: gap of unknown length
* 6417: contig of 1058 bp in length
* 7476: gap of unknown length
* 7576: contig of 1733 bp in length
* 9309: gap of unknown length
* 9409: contig of 1050 bp in length
* 10458: gap of unknown length
* 10559: contig of 1200 bp in length
* 11759: gap of unknown length
* 11858: contig of 1046 bp in length
* 12905: gap of unknown length
* 13005: contig of 1070 bp in length
* 14075: gap of unknown length
* 14174: gap of unknown length
```

```

* 14175 15258: contig of 1084 bp in length
* 15259 15358: gap of unknown length
* 15359 16585: contig of 1227 bp in length
* 16586 16885: gap of unknown length
* 16886 18061: contig of 1376 bp in length
* 18062 18161: gap of unknown length
* 18162 19504: contig of 1343 bp in length
* 19505 19604: gap of unknown length
* 19605 21161: contig of 1557 bp in length
* 21162 21261: gap of unknown length
* 21262 22375: contig of 1114 bp in length
* 22376 22475: gap of unknown length
* 22476 23942: contig of 1467 bp in length
* 23943 24042: gap of unknown length
* 24043 25894: contig of 1852 bp in length
* 25895 25994: gap of unknown length
* 25995 27573: contig of 1579 bp in length
* 27574 27673: gap of unknown length
* 27674 29345: contig of 1672 bp in length
* 29346 29445: gap of unknown length
* 29446 32419: contig of 2974 bp in length
* 32420 32519: gap of unknown length
* 32520 34592: contig of 2073 bp in length
* 34593 34692: gap of unknown length
* 34693 38815: contig of 4123 bp in length
* 38816 38915: gap of unknown length
* 38916 44847: contig of 5932 bp in length
* 44848 44947: gap of unknown length
* 44948 50756: contig of 5809 bp in length
* 50757 50856: gap of unknown length
* 50857 50959: contig of 8113 bp in length
* 50960 50959: gap of unknown length
* 50970 72750: contig of 13681 bp in length
* 72751 72850: gap of unknown length
* 72851 86053: contig of 13203 bp in length
* 86054 86153: gap of unknown length
* 86154 102699: contig of 16546 bp in length
* 102700 102799: gap of unknown length
* 102800 122169: contig of 19370 bp in length
* 122170 122269: gap of unknown length
* 122270 142795: contig of 20526 bp in length
* 142796 142895: gap of unknown length
* 142896 164429: contig of 21534 bp in length
* 164430 164539: gap of unknown length
* 164539 196146: contig of 31617 bp in length.

```

FEATURES

source

```

1. .196146
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="CT7-405C9"
/clone_lib="CitbcJ7 mouse BAC library"
BASE COUNT 52111 a 44572 c 44684 g 51449 t 3330 others
ORIGIN

```

```

Query Match 3.8%; Score 45.4; DB 2; Length 196146;
Best Local Similarity 63.1%; Pred. No. 0.028;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAAAATAAAATTCGCCATCCAGCGCTCCAGCTGAAAGTAGGCC 70
DB 3742 TAGAGAAAAACCAAGATAGACAGTGCACCAACCGAGGCTTCCAGCTAAAGAGATT 3683
QY 71 TGTCTGTCCGGTATTAAATGATGACCGTCCCGTATTTAAACAATGT 121
DB 3682 TAGCCCTGGACTTCTGTGGAATGATAGTAAAGACCTGTCTTGAAAAAATAT 3632

```

RESULT 47

AC078885/c

LOCUS

```

DEFINITION Mus musculus chromosome 6 clone RP23-117123 strain C57BL6/J,
WORKING DRAFT SEQUENCE, 44 unordered pieces.
AC078885

```

AC078885

VERSION

AC078885.8 GI:18376842

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

Mus musculus.

ORGANISM

Mus musculus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 254993)
Montgomery,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
High Throughput Mouse Sequencing
Unpublished

REFERENCE

AUTHORS

2 (bases 1 to 254993)
Montgomery,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Direct Submission

TITLE

JOURNAL

Submitted (08-AUG-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA

COMMENT

On Jan 26, 2002 this sequence version replaced gi:14488282.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpcgg.org/Sequence/mouse.html>

Contact: hpcgcmendel.mgh.harvard.edu

-----Summary Statistics

Center project name: AAF

Sequencing vector: pUC18; L08752

Chemistry: Dye-terminator Big Dye; 100%

*Consensus quality: 247096 at least Q20

*Consensus quality: 242518 at least Q30

*Consensus quality: 235589 at least Q40

Estimated insert size: agarose-FP - N/A

**Estimated insert size: 254133 - sum-of-contigs

Quality coverage: agarose-FP - N/A

Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 44 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 27415: contig of 27415 bp in length

* 27416 27435: gap of unknown length

* 27436 57817: contig of 30382 bp in length

* 57818 57837: gap of unknown length

* 57838 81637: contig of 23800 bp in length

* 81638 81657: gap of unknown length

* 81658 102155: contig of 20498 bp in length

* 102156 102175: gap of unknown length

* 102176 121200: contig of 19025 bp in length

* 121201 121221: gap of unknown length

* 121222 135073: contig of 13853 bp in length

* 135074 135093: gap of unknown length

* 135094 156689: contig of 21596 bp in length

* 156690 156710: gap of unknown length

* 156710 166186: contig of 9476 bp in length

* 166186 166205: gap of unknown length

* 166206 178818: contig of 12593 bp in length

* 178819 178799: gap of unknown length

* 178819 187247: contig of 8429 bp in length

* 187248 187267: gap of unknown length

* 187268 192492: contig of 5225 bp in length

* 192493 192513: gap of unknown length

* 192513 200084: contig of 7572 bp in length

* 200085 200104: gap of unknown length

* 200105 207054: contig of 6950 bp in length

* 207055 207074: gap of unknown length

* 207075 212944: contig of 5870 bp in length

[illegible]


```
/db_xref="GI:20559853"
/translation="MWALYVILVLENSQSVNLSFLGWRTPEGPLSLFVSLAFVLGGL
GVLSGHWLRLLIGSVSGHLPVRRAK"
1167..2216
/notes="ORF_4; wzz; similar to chain length determinant
protein"
CDS
/codon_start=1
/transl_table=11
/product="wzz"
/protein_id="AAM27625.1"
/db_xref="GI:20559854"
/translation="MTETTFQVAKNEEDLIGLRGLWAERKWLIACTIILTSLVG
TYAFLSPVYEAATVPPSWSTAGNLSRAENGLTFSVNDVSVFNLRLIAEETR
NRFYVYLSLSENERGSKDRLYNLSEMEIORSKDERIVYVIEKHDPGQAA
DWGRYVYKLAESGEMQADQAREEYKSRNIOQOQDILRDAKARRDRINRLKEA
WQYAESLKLNEPLIGGQQOQSSIMEGALNTMGTGKALRAISRALEERTSDPPI
ALRLQEQAMYSTLKLQVKNKAVAFRODGVAEVPPDPVRPKGILLVIGALSGLVWG
V
MIAFIVMAVRRAK"
2303..3568
/notes="ORF_5; similar to UDP-glucose/GDP-mannose
dehydrogenase"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27626.1"
/db_xref="GI:20559855"
/translation="MDIKAAVGLGYVGLPLAVEFGKKRSVTGFDINHISRIQLQGY
DSTLEVEKEELSAHLKFTSLSELOECNFFIVTPTIDEHKQPDLTPLVKASESI
AKVLKNDIVYESTVYGATEEYCVPLERESGKRENSDFYGVYSVFNORDNLIAL
VSTIKKVTSGPTEVAEIVDSLSREITITAGTHKAESIKVAEAKVIBNTORDNLIAL
NELAIFNKLEIDTESVLOAAGTWNELPFRPLGVLGHCIGDYPYILTHKAQSIGHP
EITLAGRLNDGAGYVYVQLVKAMLRKHHDGAVRLVMGLTFKENCPLDRLNTKYVD
IVRELAENIADVDPWVSVEEAQHEYGTLPTISAPVEGNYDAVVLVAHAHEFKELGA
DKIRAFGLSVIYDLKFLDKSDSLRL"
3578..4603
/notes="ORF_6; similar to NAD dependent
epimerase/dehydratase family"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27627.1"
/db_xref="GI:20559856"
/translation="MNAVAKALSLETTSSWLVTGVAGFIGSNLELLKLFNOKVVG
LDDSTGHYNLDEVRLGVGERWNEKLIQODRIDCDRAVEGVDDVYLHQAALGS
VPRADPTTNSVNSGLNMLVASKRAGVKSITYAASSTYGDHPLKVERVIGK
PLSPYATKYVNELVADYFSGNLFGLIRYFNVFKRQDPNGAYAAVFPKWLSAMI
KGEPTVINGDSRDFCINNVKANLLAFCETEAKNQVYNVAYGGRTTIGQLFDP
SLKVELSRNGINYEQPIYREFRAGDVRSQAETKKAQAYLGYNPEYDIQAGIAEMP
WTSFVK"
4665..5954
/notes="ORF_7; similar to Polysaccharide biosynthesis
protein; potential multiple membrane spanning domains."
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27628.1"
/db_xref="GI:20559857"
/translation="MNPKNIFNLGILGLPIIALVPHLIATVGVRFGLFALSNG
LIGAGILDLGTRAVTOKLSIRGSDSEMSYTIMSTAIRTAIVGGLGSLIIFG
LIGGFEFFSHKDVSSLEFVLLAFAPQAISATYRGVNEAYLNPKGINLRIPL
GATNFGGFLVGFYAKDLHYLTIVLSRALAFFIFRLAHGVKLETVERSKYDRR
QAVELFRFCGWTVSSIINPELVQSDREFIGVLLSAAAVTSYVIPYEITQSMILVGA
VSTVAPPSISNLIRTSFAALECFKNLILRVLLIMGGMGLAFLLPFIKLWYGDYI
GDDISVGRILCIGVFENALGAMFYSELHANVKETAILHSIELPIFILILILIPR
HGIVGNVANSURTADITIVLAVLSYFSWRVKFANC"
6001..7175
/pseudo
CDS
/codon_start=1
/transl_table=11
/1156..8091
/notes="ORF_9; similar to Glycosyl transferase"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27629.1"
/db_xref="GI:20559858"
/translation="MKSPIIDVWVWNNAGLQLRSCVDSIFKNSFGLIRKVIDNGS
DBNSCALLGERENLEVLVTGNTLGFATACNLGASKGEAEYILFLNPDALYEETIGSV
VFNMRNSNSVGVCGVQLIDIEQGVSRSCARFPSSARMLMHTMGVDRVFPALGYAMT
DWHLSSRPDVHVGAFYFRRHVKLVGGDEFERFVYLEDLDSISCSGSSFYL
STAOAFHAGGASRQIKARRFLYSRLRIYIFKNLPVWSAFILLSSLPFEPVIRSM
FSMWKLSIPALLRWSAYLMLYRWLPKWIFFNATR"
8097..9167
/notes="ORF_10; similar to Glycosyl transferases group 1"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27630.1"
/db_xref="GI:20559859"
/translation="MKILLISKYPLGASSRLTQYIDYLESKGRVDVRSFFDDGY
LSGLYSSRRSCNWLKSYISRFVFTVYRDLNWKELFFPFPVSAERFLREFKG
FVYVDDAIFHNTDLSRNLWKLKIDSVMRHSCVITAGNDILAEARVAGSSR
VBIITFDLKLRYEVRDSCFKKTVIGWISPTQKYLDDIAPLVKLCVYQASMLL
VGATPEISYFSGVYVDWKAKEDEGVYIRQMTIGIMPLDGPWEKGCCKYKLIQYM
ACGVVPASPGVGNIEIVENNESGLLAEGSRGWRESLQKLLQSSSELRECGNRSSV
EKYSLOYQSRELHQTFIEVLS"
9610..10377
/notes="ORF_11; similar to Glycosyl transferases group 1"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27631.1"
/db_xref="GI:20559860"
/translation="MAKGSKASYFRSGIAWLYRRALGKKRLRVIFQNDRDALIGLG
AITEFKSVLRSGVDLSLCHPEVPESPGAVTVTLAARLLRUKGVYFVEAANILRQSG
VSAHQVLGDLDPGNPSSIEEDFLERWSEGVCELCYRQDIASFASHIVLYLSYR
EGLPKVLYEAAACGVRVVTDPVDCRDAEADRTGLLPVDRDAVALADAIQRLVESPE
LRKMGAACTLAERDFAIESIVQOHLDIYRAGLSGA"
10395..11336
/notes="ORF_12; similar to NAD dependent
epimerase/dehydratase family"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27632.1"
/db_xref="GI:20559861"
/translation="MTGATGFVGAALVESLCSSEKYRWAGYRESSPWPEGVTPLEFL
GELGSSVAMETESVIDTVVHCAARVHMTESDPLAEFRKANVLGTLDLARQATSLG
VRRFTYLSIKYNGEGEGEQDYVADSLPNVDVPGVSKLEAQALGLAAATGLEVY
IIRPVLVYGGVKNVQTMRWLKRGPVPLPLCAIHNRSLVALDNLDTLAIQCIDHPA
ANQVFLSDGEDLSTDLLRRNGNALRTPARLPVPPVASWIGVLVAKTLNKQAFQRLC
GSLQVDIERNNRMLGWTTPGVDEAKTARVFLDHQ"
11333..12373
/notes="ORF_13; similar to Glycosyl transferase group 4;
potential multiple membrane spanning domains."
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27633.1"
/db_xref="GI:20559862"
/translation="MMSLWLLLPVAVAGLSLLTAGLRRYAIARSLMDVPNARSSHQVP
TPRGGVAVIVLSFLAVLAAFLGMDISLAIGILGAGVTALLGLDDBGHIAARWR
LICHFAGACWLLYGLGALPVPFGFMIDLVGWHIGVAVYLVWMLNLFNMDIDGI
ASVEAVCVCGAVLLVYVSGVSDASOGIWLISALLASAVTGLFWNFPARLFMGDA
GSGFLIILAGLSLQAANWSPQDFGWLLILGVVFDATLTLRLLRLLRGDKYEAHRS
HAYQASRYTGRHLPVTLAVGCINLVWLLPFLAVASGRIDGMALLIGLTLPLVAL
RFRAGTLEPRVA"
12571..>13413
/notes="ORF_14; similar to N terminal coding region of
Polysaccharide biosynthesis protein wpbm"
CDS
/codon_start=1
/transl_table=11
/protein_id="AAM27634.1"
/db_xref="GI:20559863"
/translation="MLDNLVRLLGLPRKRLQVATDIALVWLSLWLAFLVRLGTE
DMISFGGHAWLFAIPLVAVLPLIFRFGMYRAVRVYLGNDALIAVTSIALVLSL
LVWYRSPVAVVPRSLVFNFWLWSLLIGLRLAMRQYMDYSAVQSPFLNRQDGG
LPRVAYIYGAGAGNQLVALRLGRAMPVAFIDDDKOIARNYIAGLVYTAHRIQMI
DETGAQVELLAIPSATRRARREITLESLEPFLHVRSPMGFMDLASGRVKVDDIQEVDI
ADLIG"
2835 a 2899 c 3737 g 3942 t
BASE COUNT
ORIGIN
3.8%; Score 44.4; DB 1; Length 13413;
Query Match
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 15:59:59 ; Search time 2683.55 Seconds
(without alignments)
16148.066 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagtcgagatgaaag.....ctggaaggagcctggtgcac 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

- GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.pro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	1489	100.0	1489	6	AX011297	AX011297 Sequence	
2	1476.4	99.2	92077	1	AF074613	AF074613 Escherich	
3	1476.4	99.2	92077	6	AX191727	AX191727 Sequence	
4	1476.4	99.2	92721	1	AB011549	AB011549 Escherich	
5	1476.4	99.2	92721	6	AX191725	AX191725 Sequence	
6	1081.4	72.6	2407	1	ECHECCTP	X89017 E.coli 0157	
c	7	536.6	36.0	12268	1	AE013690	AE013690 Yersinia
	8	536.6	36.0	13632	1	AF135170	AF135170 Yersinia
	9	536.6	36.0	220050	1	AJ414156	AJ414156 Yersinia
10	384.8	25.8	2587	1	AB017595	AB017595 Legionell	
11	384.8	25.8	5249	1	AF276752	AF276752 Legionell	
c	12	368	24.7	1829	1	ECIS91TP	X17114 Escherichia
	13	357.6	24.0	2238	6	AR098264	AR098264 Sequence
14	349.4	23.5	11320	1	AE010762	AE010762 Methanosa	
15	348.8	23.4	17986	1	AF294823	AF294823 Shigella	
c	16	334.6	22.5	302150	1	AP001510	AP001510 Bacillus
	17	333.6	22.4	3289	1	AF486647	AF486647 Rhizobium
c	18	331.8	22.3	10578	1	AE007314	AE007314 Sinorhizo
	19	326.2	21.9	4754	1	RLE438039	AJ438039 Rhizobium
20	323.2	21.7	340857	1	AP003010	AP003010 Mesorhizo	
c	21	314.8	21.1	198050	1	AL646061	AL646061 Ralstonia
	22	309.2	20.8	2904	1	SRFRPS	Y14317 Streptomyce
23	305.4	20.5	2805	1	ECOKATGA	M21516 E.coli katG	
24	305.4	20.5	13840	1	AE000468	AE000468 Escherich	
25	305.4	20.5	96484	1	ECOUW87	L19201 E. coli chr	
26	303.8	20.4	318703	1	AP002567	AP002567 Escherich	
27	303.4	20.4	3018	8	AN1305225	AJ305225 Aspergill	
28	303	20.3	10667	1	AE005625	AE005625 Escherich	
29	302.2	20.3	2885	1	AF027168	AF027168 Caulobact	
30	302.2	20.3	10677	1	AE005967	AE005967 Caulobact	
31	300.6	20.2	2487	1	HMY16851	Y16851 Haloarcula	
32	298.4	20.0	2223	1	AF314111	AF314111 Mycobact	
33	298.4	20.0	2223	1	AF314115	AF314115 Mycobact	
34	298.4	20.0	2223	1	AF314117	AF314117 Mycobact	
35	298.4	20.0	2223	1	AF314120	AF314120 Mycobact	
36	298.4	20.0	2223	1	AF314121	AF314121 Mycobact	
37	298.4	20.0	2223	1	MTU40593	U40593 Mycobacter	
38	298.4	20.0	2223	1	MTU40595	U40595 Mycobacter	
39	298.4	20.0	2223	1	MTU41305	U41305 Mycobacter	
40	298.4	20.0	2223	1	MTU41306	U41306 Mycobacter	
41	298.4	20.0	2223	1	MTU41310	U41310 Mycobacter	
42	298.4	20.0	2223	1	MTU41313	U41313 Mycobacter	
43	298.4	20.0	2223	1	MTU41314	U41314 Mycobacter	
44	298.4	20.0	2225	1	MTU41304	U41304 Mycobacter	
45	298.4	20.0	2235	6	I61339	I61339 Sequence 1	
46	298.4	20.0	2235	6	I74293	I74293 Sequence 1	
47	298.4	20.0	2297	1	MTU06289	U06289 Mycobacter	
48	298.4	20.0	2311	1	MTU06258	U06258 Mycobacter	
49	298.4	20.0	2316	1	MTU06264	U06264 Mycobacter	
50	298.4	20.0	2326	1	MTU06260	U06260 Mycobacter	
51	298.4	20.0	2331	6	I61357	I61357 Sequence 20	
c	52	298.4	20.0	16760	1	AE007051	AE007051 Mycobact
	53	298.4	20.0	44201	1	MTCY180	297193 Mycobacter
c	54	298.2	20.0	2359	1	MTU06262	U06262 Mycobacter
	55	298	20.0	2322	1	MTU06267	U06267 Mycobacter
56	297.4	20.0	2326	1	MTU06270	U06270 Mycobacter	
57	297.2	20.0	3027	1	AF126956	AF126956 Streptom	
58	297.2	20.0	10230	1	AE000951	AE000951 Archaeogl	
59	297	19.9	2208	1	AB020234	AB020234 Bacillus	
60	297	19.9	2232	1	AB020065	AB020065 Bacillus	
61	297	19.9	2232	1	AB020120	AB020120 Bacillus	
62	297	19.9	2235	1	AB020067	AB020067 Bacillus	
63	297	19.9	2235	1	AB020092	AB020092 Bacillus	
64	297	19.9	2238	1	AB020115	AB020115 Bacillus	
65	297	19.9	2241	1	AB020094	AB020094 Bacillus	

66	297	19.9	2247	1	AB020078	AB020078	Bacillus	139	296.8	19.9	4810	1	MTKATGCP	X58081	M. tuberculosis
67	297	19.9	2250	1	AB020106	AB020106	Bacillus	140	296.4	19.9	2308	1	MTU06266	U06266	Mycobacteri
68	297	19.9	2253	1	AB020116	AB020116	Bacillus	141	296.4	19.9	2324	1	MTU06268	U06268	Mycobacteri
69	297	19.9	2253	1	AB020117	AB020117	Bacillus	142	295.8	19.9	2316	1	MTU06265	U06265	Mycobacteri
70	297	19.9	2256	1	AB020064	AB020064	Bacillus	143	295.6	19.9	14922	1	SCF73	AL121746	Streptomy
71	297	19.9	2256	1	AB020066	AB020066	Bacillus	144	295	19.8	2312	1	MTU06263	U06263	Mycobacteri
72	297	19.9	2256	1	AB020071	AB020071	Bacillus	145	294.8	19.8	2332	1	MTU06272	U06272	Mycobacteri
73	297	19.9	2256	1	AB020072	AB020072	Bacillus	146	293.8	19.7	2208	1	AB020235	AB020235	Bacillus
74	297	19.9	2256	1	AB020073	AB020073	Bacillus	147	293.8	19.7	2238	1	AB020123	AB020123	Bacillus
75	297	19.9	2256	1	AB020074	AB020074	Bacillus	148	293.8	19.7	2238	1	AB020156	AB020156	Bacillus
76	297	19.9	2256	1	AB020075	AB020075	Bacillus	149	293.8	19.7	2244	1	AB020158	AB020158	Bacillus
77	297	19.9	2256	1	AB020076	AB020076	Bacillus	150	293.8	19.7	2256	1	AB020122	AB020122	Bacillus
78	297	19.9	2256	1	AB020077	AB020077	Bacillus								
79	297	19.9	2256	1	AB020079	AB020079	Bacillus								
80	297	19.9	2256	1	AB020080	AB020080	Bacillus								
81	297	19.9	2256	1	AB020081	AB020081	Bacillus								
82	297	19.9	2256	1	AB020082	AB020082	Bacillus								
83	297	19.9	2256	1	AB020083	AB020083	Bacillus								
84	297	19.9	2256	1	AB020084	AB020084	Bacillus								
85	297	19.9	2256	1	AB020085	AB020085	Bacillus								
86	297	19.9	2256	1	AB020086	AB020086	Bacillus								
87	297	19.9	2256	1	AB020087	AB020087	Bacillus								
88	297	19.9	2256	1	AB020088	AB020088	Bacillus								
89	297	19.9	2256	1	AB020090	AB020090	Bacillus								
90	297	19.9	2256	1	AB020091	AB020091	Bacillus								
91	297	19.9	2256	1	AB020093	AB020093	Bacillus								
92	297	19.9	2256	1	AB020096	AB020096	Bacillus								
93	297	19.9	2256	1	AB020097	AB020097	Bacillus								
94	297	19.9	2256	1	AB020098	AB020098	Bacillus								
95	297	19.9	2256	1	AB020099	AB020099	Bacillus								
96	297	19.9	2256	1	AB020101	AB020101	Bacillus								
97	297	19.9	2256	1	AB020102	AB020102	Bacillus								
98	297	19.9	2256	1	AB020104	AB020104	Bacillus								
99	297	19.9	2256	1	AB020105	AB020105	Bacillus								
100	297	19.9	2256	1	AB020107	AB020107	Bacillus								
101	297	19.9	2256	1	AB020108	AB020108	Bacillus								
102	297	19.9	2256	1	AB020109	AB020109	Bacillus								
103	297	19.9	2256	1	AB020110	AB020110	Bacillus								
104	297	19.9	2256	1	AB020111	AB020111	Bacillus								
105	297	19.9	2256	1	AB020112	AB020112	Bacillus								
106	297	19.9	2256	1	AB020113	AB020113	Bacillus								
107	297	19.9	2256	1	AB020114	AB020114	Bacillus								
108	297	19.9	2256	1	AB020118	AB020118	Bacillus								
109	297	19.9	2259	1	AB020068	AB020068	Bacillus								
110	297	19.9	2259	1	AB020070	AB020070	Bacillus								
111	297	19.9	2259	1	AB020089	AB020089	Bacillus								
112	297	19.9	2259	1	AB020095	AB020095	Bacillus								
113	297	19.9	2259	1	AB020100	AB020100	Bacillus								
114	297	19.9	2259	1	AB020121	AB020121	Bacillus								
115	297	19.9	2259	1	AB020069	AB020069	Bacillus								
116	297	19.9	2262	1	AB020119	AB020119	Bacillus								
117	297	19.9	2265	1	AB020103	AB020103	Bacillus								
118	297	19.9	2642	1	BACPERA	M29876	B. stearother								
119	296.8	19.9	2223	1	AF314107	AF314107	Mycobacte								
120	296.8	19.9	2223	1	AF314108	AF314108	Mycobacte								
121	296.8	19.9	2223	1	AF314109	AF314109	Mycobacte								
122	296.8	19.9	2223	1	AF314110	AF314110	Mycobacte								
123	296.8	19.9	2223	1	AF314112	AF314112	Mycobacte								
124	296.8	19.9	2223	1	AF314113	AF314113	Mycobacte								
125	296.8	19.9	2223	1	AF314114	AF314114	Mycobacte								
126	296.8	19.9	2223	1	AF314116	AF314116	Mycobacte								
127	296.8	19.9	2223	1	AF314118	AF314118	Mycobacte								
128	296.8	19.9	2223	1	AF314119	AF314119	Mycobacte								
129	296.8	19.9	2223	1	AF314122	AF314122	Mycobacte								
130	296.8	19.9	2223	1	AF314124	AF314124	Mycobacte								
131	296.8	19.9	2223	1	MBDNAKATG	X83277	M. bovis kat								
132	296.8	19.9	2223	1	MTU41307	U41307	Mycobacteri								
133	296.8	19.9	2223	1	MTU41308	U41308	Mycobacteri								
134	296.8	19.9	2223	1	MTU41309	U41309	Mycobacteri								
135	296.8	19.9	2223	1	MTU41311	U41311	Mycobacteri								
136	296.8	19.9	2223	1	MTU41312	U41312	Mycobacteri								
137	296.8	19.9	2224	1	AF314123	AF314123	Mycobacte								
138	296.8	19.9	2339	1	MTU06259	U06259	Mycobacteri								

ALIGNMENTS

RESULT 1

AX011297

LOCUS

AX011297

DEFINITION

Sequence 1 from Patent WO9955908.

ACCESSION

AX011297

VERSION

AX011297.1 GI:9997847

KEYWORDS

Escherichia coli.

SOURCE

Escherichia coli.

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE

1 (bases 1 to 1489)

AUTHORS

Thierry D., Frechon, D.T. and Laure, F.C.

TITLE

Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)

JOURNAL

Patent: WO 9955908-A 1 04-NOV-1999;

THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES

Location/Qualifiers

1..1489

source

/organism="Escherichia coli"

/db_xref="taxon:562"

BASE COUNT

386 a 354 c 391 g 358 t

ORIGIN

Query Match 100.0%; Score 1489; DB 6; Length 1489;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGTCCGGAGATGAAGCACCACCTGTGTACCCCATCAGCGTGTCCCGCAGGCCA 60

|||||

Db 1 CTGAGTCCGGAGATGAAGCACCACCTGTGTACCCCATCAGCGTGTCCCGCAGGCCA 60

|||||

QY 61 TGATTTTGTACAGACTCAATGACTCCGAGGAGTGAACCTTCGGTTGTTCTCCA 120

|||||

Db 61 TGATTTTGTACAGACTCAATGACTCCGAGGAGTGAACCTTCGGTTGTTCTCCA 120

|||||

QY 121 GCAGTTAAGCAGCGGTTCCCTGCTCAAAAATGTCGCAAAACGGGGAAGCATCAGAA 180

|||||

Db 121 GCAGTTAAGCAGCGGTTCCCTGCTCAAAAATGTCGCAAAACGGGGAAGCATCAGAA 180

|||||

QY 181 GGGCGGGGGAATCCGTCGCGGCAAGTGAACCGTGCACACTCCGGGAGTAGTACATCGCGCC 240

|||||

Db 181 GGGCGGGGGAATCCGTCGCGGCAAGTGAACCGTGCACACTCCGGGAGTAGTACATCGCGCC 240

|||||

QY 241 GCGCGTGTATACGGGCAAGATGGTGGCAAACTCCGCTCCGTCGAGCGGGGTATTTCAGG 300

|||||

Db 241 GCGCGTGTATACGGGCAAGATGGTGGCAAACTCCGCTCCGTCGAGCGGGGTATTTCAGG 300

|||||

QY 301 ATACCTTCGTCATCAACACAGTACAAAACCAAGACACAGCTTTTGTCTTGACATCCAC 360

|||||

Db 301 ATACCTTCGTCATCAACACAGTACAAAACCAAGACACAGCTTTTGTCTTGACATCCAC 360

|||||

QY 361 AAAGAGGGAATATTTCAGTCTCGCGCAGCTCAACGCGCATCGTCAGTTGGGCTTGAA 420

|||||

Db 361 AAAGAGGGAATATTTCAGTCTCGCGCAGCTCAACGCGCATCGTCAGTTGGGCTTGAA 420

```
QY 421 CCCCTTAGTATTTTCTGCTAGTATCTATCCACCAATAGTATATCTCTGTGTCATCA 480
DB 421 CCCCTTAGTATTTTCTGCTAGTATCTATCCACCAATAGTATATCTCTGTGTCATCA 480
QY 481 ATAAAGCTTACCTTTCTGTATACAACTGCGAATTTCCCTTAATCCGGAGCTATTCGATGA 540
DB 481 ATAAAGCTTACCTTTCTGTATACAACTGCGAATTTCCCTTAATCCGGAGCTATTCGATGA 540
QY 541 TAAAAAAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 TAAAAAAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 TAGCCCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTC 660
DB 601 TAGCCCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTC 660
QY 661 CTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCGCACCA 720
DB 661 CTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCGCACCA 720
QY 721 GATTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACTT 780
DB 721 GATTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACTT 780
QY 781 CCCAGGATGGTCCCTGCGGATATGCTCATATGCTCTCTTTCTTTTATCTGATGGCTT 840
DB 781 CCCAGGATGGTCCCTGCGGATATGCTCATATGCTCTCTTTCTTTTATCTGATGGCTT 840
QY 841 GGCACGGTGGCGGAACATACAGACATATGATGGCGGGGAGCGCGAGTGGTGGTCCAG 900
DB 841 GGCACGGTGGCGGAACATACAGACATATGATGGCGGGGAGCGCGAGTGGTGGTCCAG 900
QY 901 AACGTTTTGAACCGCTGAACAGCTGGCGGATATCTCTGATATGATAAGCCCGCTCGAT 960
DB 901 AACGTTTTGAACCGCTGAACAGCTGGCGGATATCTCTGATATGATAAGCCCGCTCGAT 960
QY 961 TGCTGTGGCGAGTCAGAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGGTCC 1020
DB 961 TGCTGTGGCGAGTCAGAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGGTCC 1020
QY 1021 TGACTGGTATGTTGCTTGAATCCATGGGATTTAAAGCGCTGGGATTTGCTGGCGGAA 1080
DB 1021 TGACTGGTATGTTGCTTGAATCCATGGGATTTAAAGCGCTGGGATTTGCTGGCGGAA 1080
QY 1081 GAGAAGATGACTGGGAGTGGACCTGGTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 GAGAAGATGACTGGGAGTGGACCTGGTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 ATAAACGGGATAAAAACGGGAACTTCAGAAACCTCTGCGCCAGCAGATGGGACTTA 1200
DB 1141 ATAAACGGGATAAAAACGGGAACTTCAGAAACCTCTGCGCCAGCAGATGGGACTTA 1200
QY 1201 TTTATGTCATCTGAAAGCCCGGCTGGAAGACAGATCTCTGCTGCTGCTGCTGCTGCTG 1260
DB 1201 TTTATGTCATCTGAAAGCCCGGCTGGAAGACAGATCTCTGCTGCTGCTGCTGCTGCTG 1260
QY 1261 TCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGCTGTGGCCCTGATCGCG 1320
DB 1261 TCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGCTGTGGCCCTGATCGCG 1320
QY 1321 GAGGGATACATTTGTTAAGACATATGTCAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 1321 GAGGGATACATTTGTTAAGACATATGTCAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1381 GGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1381 GGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 GAAACGGCAATATACCATCACCAGTGGCTTGAAGAGAGCTGCTGCTGCTGCTGCTGCTG 1489
DB 1441 GAAACGGCAATATACCATCACCAGTGGCTTGAAGAGAGCTGCTGCTGCTGCTGCTGCTG 1489
```

```
RESULT 2
AF074613 92077 bp DNA circular BCT 04-NOV-1998
LOCUS Escherichia coli O157:H7 plasmid pO157, complete sequence.
DEFINITION AF074613
ACCESSION AF074613
VERSION AF074613.1 GI:3822114
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 92077)
Burland,V., Shao,Y., Perna,N.T., Plunkett,G., Sofia,H.J. and
Blattner,F.R.
TITLE The complete DNA sequence and analysis of the large virulence
plasmid of Escherichia coli O157:H7
JOURNAL Nucleic Acids Res. 26 (18), 4196-4204 (1998)
MEDLINE 98391744
PUBMED 9722640
REFERENCE 2 (bases 1 to 92077)
Burland,V., Shao,Y., Perna,N.T., Plunkett,G. III, Sofia,H.J. and
Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. .92077
/organism="Escherichia coli O157:H7"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:83334"
/lab_host="Escherichia coli C600"
/plasmid="pO157"
1. .561
/gene="fno"
/gene="fno"
/note="L7001"
1. .561
/gene="fno"
/note="97 pct identical amino acid sequence and equal
length to F101_ECOLI SW: P22707"
/codon_start=1
/transl_table=11
/product="fertility inhibition protein (conjugal transfer
repressor)"
/protein_id="AAC70069.1"
/db_xref="GI:3822115"
/translation="MAEOKRPVLTARKTEGETPVRSRKTIINVTTPKWKVKOKLA
EKRAAEALAAKAAQARQALSIYLNLPOTDEAVNTLKPWPGLFDGTPRLACGIND
VLEEDVAQRNIPULSHKRLRAKAITRSESYLCAMKAGACRYDTEGYVTEHISQEBEA
YAAARLDKTRQRNRIKAEQLQAVLDEK"
653. .681
/promoter
/note="predicted sigma 70 promoter; score of 52%"
698. .949
/gene="L7002"
698. .949
/gene="L7002"
/note="54 pct identical (1 gap) to 23 residues of an
approx. 269 aa protein EXOK_RHIME SW: P33693"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC70070.1"
/db_xref="GI:3822116"
/translation="MDSEVHGTVRSGVTSVPEAGPLFWKSVDAWKQRKHGDLPLV
HPLTGGSSLPKRGNTATGAAEGGNEKSLHYSRDEG"
1025. .1052
/promoter
/note="predicted sigma 70 promoter; score of 62%"
1151. .1612
/gene="L7003"
1151. .1612
/gene="L7003"
/note="98 pct identical and equal length to YF13_ECOLI SW:
```

```

Q99342"
/codon_start=1
/transl_table=11
/product="hypothetical protein 15.6 kDa protein in finO 3'
region precursor"
/protein_id="AAC70071.1"
/db_xref="GI:3822117"
/transl_table=11
/translation="MRKIPLVLFISWFLCADIHGRVVRVLDGDTIEVMSRKAVR
IRLVNIDAPEKQDYGWSDMMKSLVAGKTVTVTFORDRYGRMLGQVYADGMNVN
QFVVRAGAANVYEQYNTDPLVPLVQNEARQKRGLWSDADPPVPIWRHK"
1390..1419
/notes="predicted sigma 70 promoter; score of 56%"
/gene="L7004"
1638..1867
/gene="L7004"
1658..1867
/gene="L7004"
/notes="55 pct identical (0 gaps) to 66 residues of an
approx. 72 aa protein HHA_ECOLI SW: P23870"
/codon_start=1
/transl_table=11
/product="putative hemolysin expression modulating
protein"
/protein_id="AAC70072.1"
/db_xref="GI:3822118"
/translation="MEKTKQEWLYQLRSCSVNTLEKIIHKNRDSLSTSERSEFNSAA
DHKLAEITGKLYDRIPKEIKYVR"
1899..1927
/notes="predicted sigma 70 promoter; score of 56%"
/gene="L7005"
1905..2243
/gene="L7005"
1905..2243
/gene="L7005"
/notes="98 pct identical (0 gaps) to 57 residues of an
approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF
4, TRSW: Q52340"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC70073.1"
/db_xref="GI:3822119"
/translation="MKLIIFILVLIIAALLIRILRSVNOHSPLIMQLHAAGIRTGD
AERILSGEYDASHLRPERRRDILLEVLQAGIPLLRSHDARKLQWGEWLTIG
AAQSPHRS"
complement(2240..2327)
/notes="initiation site; Escherichia coli plasmid R100 ssIB
gene; 99 pct identical to ssIB locus R108SIB accession
D90185"
2416..2421
/notes="100 pct identical (0 gaps) to the -35 region at
23..28 locus ECNR1REP accession X02302"
2439..2445
/notes="100 pct identical (0 gaps) to the -10 region at
46..52 locus ECNR1REP accession X02302"
2473..2478
/notes="100 pct identical (0 gaps) to RBS at 80..85 locus
ECNR1REP accession X02302"
2483..2737
/gene="cpb2"
/notes="L7006"
2483..2737
/gene="cpb2"
/notes="98 pct identical amino acid sequence and equal
length to CPB2_ECOLI SW: P03847"
/codon_start=1
/transl_table=11
/product="CcpB protein (RepA2 protein)"
/protein_id="AAC70074.1"
/db_xref="GI:3822120"
/translation="MSQTENAVTSSSGAKRAYRKGNPLSDAEKQLSLVARKRSPKEY
KVLEPKYKAMLMQMCHEDGLTQAEVLTALIKSEAQRKV"
2754..2782
/notes="predicted sigma 70 promoter; score of 69%"
2973..3047

```

CDS
 2973..3047
 /gene="L7007"
 /note="91 pct identical (0 gaps) to 24 residues of an
 approx. 24 aa protein uORF P30REFFIC, accession M16167,
 translationally coupled to replication initiation protein"
 /codon_start=1
 /transl_table=11
 /product="replication initiation protein"
 /protein_id="AAC70075.1"
 /db_xref="GI:3822121"
 /translation="MLGRVQDFFLCSLLLRIVSAGWCD"
 3032..3035
 /note="100 pct identical (0 gaps) to RBS at 641..644 locus
 ECNR1REP accession X02302"
 3040..3897
 /gene="rep2"
 /note="L7008"
 3040..3897
 /gene="rep2"
 /note="96 pct identical amino acid sequence and equal
 length to REP2_ECOLI SW: P03066"
 /codon_start=1
 /transl_table=11
 /product="replication initiation protein"
 /protein_id="AAC70076.1"
 /db_xref="GI:3822122"
 /translation="MTDLQOYYRQVKNPNPVPFREGAGTLKFCEKLMKAVGFTSR
 FDPATHVAHRSKGLRRPMPVLRRAIDALLOGLCFHYDPLANRVQCSITTLAIEGG
 LATESNAGKLSITRATRALFELSELGLITYOTEYDPLIGCVIPDITFTFALFALDV
 SEDAVAAARRSRQVEMLRKKQGLDILGMDELIAKAWRFVRFRFSTQYTELKSRGK
 RARRRAGRRQDITVLVKRQLTREISEGRFTANREAVKREVRVERKERMILSRNRN
 YSLRATSP"
 3847..4095
 /standard_name="CIS"
 /note="88 pct identical to locus ECCIS accession X12587,
 required for cis-activation of orir by the replication
 initiation protein"
 4072..4080
 /note="dnaA site; 100 pct identical (0 gaps) to locus
 ECNR1REP at (1682..1690) accession X02302"
 4084..4232
 /standard_name="orir"
 /note="89 pct identical to orir (1094..1242); minimum
 segment for replication of E. coli IncFII plasmid NR1
 ECRRPAL X12776"
 /direction=right
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

Query Match 99.2%; Score 1476.4; DB 1; Length 92077;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCAGTCCGGAGATGAA-AGCACACATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
 |||||
 Db 6836 CTGCAGTCCGGAGATGAACAGCACCATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 6895
 |||||
 QY 60 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 119
 |||||
 Db 6896 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 6955
 |||||
 QY 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 179
 |||||
 Db 6956 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 7015

misc_feature
 4072..4080
 /note="dnaA site; 100 pct identical (0 gaps) to locus
 ECNR1REP at (1682..1690) accession X02302"
 4084..4232
 /standard_name="orir"
 /note="89 pct identical to orir (1094..1242); minimum
 segment for replication of E. coli IncFII plasmid NR1
 ECRRPAL X12776"
 /direction=right
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

promoter
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

gene
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

CDS
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

Query Match 99.2%; Score 1476.4; DB 1; Length 92077;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCAGTCCGGAGATGAA-AGCACACATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
 |||||
 Db 6836 CTGCAGTCCGGAGATGAACAGCACCATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 6895
 |||||
 QY 60 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 119
 |||||
 Db 6896 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 6955
 |||||
 QY 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 179
 |||||
 Db 6956 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 7015

misc_feature
 4072..4080
 /note="dnaA site; 100 pct identical (0 gaps) to locus
 ECNR1REP at (1682..1690) accession X02302"
 4084..4232
 /standard_name="orir"
 /note="89 pct identical to orir (1094..1242); minimum
 segment for replication of E. coli IncFII plasmid NR1
 ECRRPAL X12776"
 /direction=right
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

promoter
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

gene
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

CDS
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

Query Match 99.2%; Score 1476.4; DB 1; Length 92077;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCAGTCCGGAGATGAA-AGCACACATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
 |||||
 Db 6836 CTGCAGTCCGGAGATGAACAGCACCATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 6895
 |||||
 QY 60 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 119
 |||||
 Db 6896 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 6955
 |||||
 QY 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 179
 |||||
 Db 6956 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 7015

misc_feature
 4072..4080
 /note="dnaA site; 100 pct identical (0 gaps) to locus
 ECNR1REP at (1682..1690) accession X02302"
 4084..4232
 /standard_name="orir"
 /note="89 pct identical to orir (1094..1242); minimum
 segment for replication of E. coli IncFII plasmid NR1
 ECRRPAL X12776"
 /direction=right
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

promoter
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

gene
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

CDS
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

Query Match 99.2%; Score 1476.4; DB 1; Length 92077;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCAGTCCGGAGATGAA-AGCACACATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
 |||||
 Db 6836 CTGCAGTCCGGAGATGAACAGCACCATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 6895
 |||||
 QY 60 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 119
 |||||
 Db 6896 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 6955
 |||||
 QY 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 179
 |||||
 Db 6956 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 7015

misc_feature
 4072..4080
 /note="dnaA site; 100 pct identical (0 gaps) to locus
 ECNR1REP at (1682..1690) accession X02302"
 4084..4232
 /standard_name="orir"
 /note="89 pct identical to orir (1094..1242); minimum
 segment for replication of E. coli IncFII plasmid NR1
 ECRRPAL X12776"
 /direction=right
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

promoter
 4169..4197
 /note="predicted sigma 70 promoter; score of 56%"
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

gene
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

CDS
 4259..4453
 /gene="L7009"
 4259..4453
 /gene="L7009"
 /note="95 pct identical to (0 gaps) 64 residues of a 128
 aa protein REPA4 locus ECRC1 accession V00351"
 /codon_start=1
 /transl_table=11

Query Match 99.2%; Score 1476.4; DB 1; Length 92077;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCAGTCCGGAGATGAA-AGCACACATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
 |||||
 Db 6836 CTGCAGTCCGGAGATGAACAGCACCATGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 6895
 |||||
 QY 60 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 119
 |||||
 Db 6896 ATGATTTTGTACAGACTCAATGACTACCGGAGCGACTGAACCTTCGGTGTCTTC 6955
 |||||
 QY 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 179
 |||||
 Db 6956 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGCGGAAAAACGCGGGAAGCATCAGA 7015

QY	180	AGGCGGGGNACTCCGTCGGGCCAGTGAACCGTGCACACTCGGGCACTACATGCCGC	239
DB	7016	AGGCGGGGGAACCTCCGTCGGGCCAGTGAACCGTGCACACTCGGGCACTACATGCCGC	7075
QY	240	CGGCGCTGATACCGGCAAGAATGGTGCACAACTCCGCTCCGTGCAGCGGCGCTATTTTCAG	299
DB	7076	CGGCGCTGATACCGGCAAGAATGGTGCACAACTCCGCTCCGTGCAGCGGCGCTATTTTCAG	7135
QY	300	GATACCCCTTCGTCAATCAACACGTACAAACAGAGACACAGCTTTTGTCTTGACATCCA	359
DB	7136	GATACCCCTTCGTCAATCAACACGTACAAACAGAGACACAGCTTTTGTCTTGACATCCA	7195
QY	360	CAAGAAGGGAATATTACGCTCTCCGAGACACTCAAGGCGATCGTCAGTTGCGCGCTTGA	419
DB	7196	CAAGAAGGGAATATTACGCTCTCCGAGACACTCAAGGCGATCGTCAGTTGCGCGCTTGA	7255
QY	420	ACCCCTTAGTATTTTGTCTGTAGTATCTATCCACGAATAGTATATCCTGTGTCATC	479
DB	7256	ACCCCTTAGTATTTTGTCTGTAGTATCTATCCACGAATAGTATATCCTGTGTCATC	7315
QY	480	AATAAGTGTGACTTTTGTATACAACTGCGAAATTCCTTAATCCGGAGCTATTCGTATG	539
DB	7316	AATAAGTGTGACTTTTGTATACAACTGCGAAATTCCTTAATCCGGAGCTATTCGTATG	7375
QY	540	ATAAAAAAACTCTTCCTGTCTGATCTCTCGCGCTATCGGGAGCTTTTCTACCGCT	599
DB	7376	ATAAAAAAACTCTTCCTGTCTGATCTCTCGCGCTATCGGGAGCTTTTCTACCGCT	7435
QY	600	GTAGCCGCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTTGGATTTAACT	659
DB	7436	GTAGCCGCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTTGGATTTAACT	7495
QY	660	CCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTGATATAGCCACC	719
DB	7496	CCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTGATATAGCCACC	7555
QY	720	AGATTTCAACAGCTGGGATATGGAGGCTCTCAAAAAAGATATCAAGATTTGCTGCACAACT	779
DB	7556	AGATTTCAACAGCTGGGATATGGAGGCTCTCAAAAAAGATATCAAGATTTGCTGCACAACT	7615
QY	780	TCCAGGATTGGTCCCTGCGGATTATGGTCATTATGGTCTCTTTTATTCGTATGGCT	839
DB	7616	TCCAGGATTGGTCCCTGCGGATTATGGTCATTATGGTCTCTTTTATTCGTATGGCT	7675
QY	840	TGGCACCGTGCCGGAACATACAGACATATATGCGCGGGAGCGCCAGTGTGTGTGTCAG	899
DB	7676	TGGCACCGTGCCGGAACATACAGACATATATGCGCGGGAGCGCCAGTGTGTGTGTCAG	7735
QY	900	CAACGTTTTGAACCGCTGAACAGCTGCGCGGATAACGTTAATCTGGATAAGCCCTGCA	959
DB	7736	CAACGTTTTGAACCGCTGAACAGCTGCGCGGATAACGTTAATCTGGATAAGCCCTGCA	7795
QY	960	TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCCCTGGGAGACCTGATGGTC	1019
DB	7796	TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCCCTGGGAGACCTGATGGTC	7855
QY	1020	CTGACTGGTAAATGTGGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA	1079
DB	7856	CTGACTGGTAAATGTGGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA	7915
QY	1080	AGAGAAGTACTGGGAGTCGGACCTCGTATACTTGGGGGCTCAGCAACAGCCCTTGTGCA	1139
DB	7916	AGAGAAGTACTGGGAGTCGGACCTCGTATCTGGGGGCTCAGCAACAGCCCTTGTGCA	7975
QY	1140	GATAACCGGATAAAAACGGGAACTTCAGAAACCTCTTCCGCCACGCGAGATGGACTT	1199
DB	7976	GATAACCGGATAAAAACGGGAACTTCAGAAACCTCTTCCGCCACGCGAGATGGACTT	8035
QY	1200	ATTATATGTCATTCCTGAAGCCCGGTGGAAAAACCGAGATCCTCTGGCTCCGCGAAGAT	1259
DB	8036	ATTATATGTCATTCCTGAAGCCCGGTGGAAAAACCGAGATCCTCTGGCTCCGCGAAGAT	8095

QY	1260	ATCAGGGAAGCTTTTTCACGTATGGCCATCGATGATGAGAGACTGTGGCCCTGATGTCGCG	1319
Db	8096	ATCAGGGAAGCTTTTTCACGTATGGCCATCGATGATGAGAGACTGTGGCCCTGATGTCGCG	8155
QY	1320	GGAGGCATACATTGTGTAAGACACATGGTCAGCGTCTCCTGAAAAATGTAATTTGGCGCA	1379
Db	8156	GGAGGCATACATTGTGTAAGACACATGGTCAGCGTCTCCTGAAAAATGTAATTTGGCGCA	8215
QY	1380	GGGCCTGATGTGCACCTGTGGAGAGCAGGACTGGGATGGAAAAATAAATTTGGTACA	1439
Db	8216	GGGCCTGATGTGCACCTGTGGAGAGCAGGACTGGGATGGAAAAATAAATTTGGTACA	8275
QY	1440	GGAACGGCAAAATATACATCACACAGTGGCCTGGAAGAGCCTGGTGCAC	1489
Db	8276	GGAACGGCAAAATATACATCACACAGTGGCCTGGAAGAGCCTGGTGCAC	8325
RESULT 3			
AX191727			
LOCUS	AX191727	92077 bp	DNA linear PAT 15-AUG-2001
DEFINITION	Sequence 9 from Patent WO0149775.		
ACCESSION	AX191727		
VERSION	AX191727.1	GI:15209896	
KEYWORDS			
SOURCE	Escherichia coli.		
ORGANISM	Escherichia coli		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
AUTHORS	Escherichia.		
TITLE	Iversen, P.L.		
JOURNAL	Antisense antibacterial cell division composition and method		
FEATURES	Patent: WO 0149775-A 9 12-JUL-2001;		
source	Avi Biopharma, Inc. (US)		
	Location/Qualifiers		
	1..92077		
	/organism="Escherichia coli"		
	/db_xref="taxon:562"		
BASE COUNT	24693 a 20133 c 23728 g 23523 t		
ORIGIN			
	Query Match 99.28; Score 1476.4; DB 6; Length 92077;		
	Best Local Similarity 99.98; Pred. No. 0;		
	Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;		
QY	1	CTGCAGTCCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC	59
Db	6836	CTGCAGTCCGGAGATGAAAGCAGCACCACCTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC	6895
QY	60	ATGATTTTTGTACAGACTCAATGACTACCGGAGCAGTGAACCTTCCGGTTGTGTTCTCC	119
Db	6896	ATGATTTTTGTACAGACTCAATGACTACCGGAGCAGTGAACCTTCCGGTTGTGTTCTCC	6955
QY	120	AGCAGTTAAGCCAGCGTTTCCCTGCTGAAAAATGTCGGCAACGGGGAAGCATCAGA	179
Db	6956	AGCAGTTAAGCCAGCGTTTCCCTGCTGAAAAATGTCGGCAACGGGGAAGCATCAGA	7015
QY	180	AGGCGGGGGAACCTCCGTCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCG	239
Db	7016	AGGCGGGGGAACCTCCGTCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCG	7075
QY	240	CGCGCCTGATACCGGCAAGAAATGGTGCACAACTCCCGCTCCGTCAGCGGGCTATTTCAG	299
Db	7076	CGCGCCTGATACCGGCAAGAAATGGTGCACAACTCCCGCTCCGTCAGCGGGCTATTTCAG	7135
QY	300	GATACCCCTTCGTATCAACAGGTACAACACAGAGACCAGCTTTTGTCTGACATCCA	359
Db	7136	GATACCCCTTCGTATCAACAGGTACAACACAGAGACCAGCTTTTGTCTGACATCCA	7195
QY	360	CAAGAAGGGAATATTCAGTCTCGCAGCAGCTCAACGGCATCGTCAGTTGGCGCTTGA	419
Db	7196	CAAGAAGGGAATATTCAGTCTCGCAGCAGCTCAACGGCATCGTCAGTTGGCGCTTGA	7255
QY	420	ACCCCTTAGTATTTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTCTGTCATC	479

|||||
7256 ACCCTTAGTATTTTCTGTAGTATCTATCCAGCAATAGTATATCTCTGTCATC 7315
QY 480 AATAAGTTGACTTTTGTATACACATCGCAATTTCCCTTAATCCGAGCTATTCGTATG 539
Db 7316 AATAAGTTGACTTTTGTATACACATCGCAATTTCCCTTAATCCGAGCTATTCGTATG 7375
QY 540 AATAAAAACTCTCTCTGTTCTGATCTCTCTGCGCTATCGGGAGCTTTTCTACCGCT 599
Db 7376 AATAAAAACTCTCTCTGTTCTGATCTCTCTGCGCTATCGGGAGCTTTTCTACCGCT 7435
QY 600 GTAGCCGCTGATAAAAAAGACATCAAAATTTCTACTATCCAGAAACACTGGATTAACT 659
Db 7436 GTAGCCGCTGATAAAAAAGACATCAAAATTTCTACTATCCAGAAACACTGGATTAACT 7495
QY 660 CTTCTGAGATTACACAGCCCTGGAATCAAAATCCCTGGGGGCTGATTTGATTATGCCACC 719
Db 7496 CTTCTGAGATTACACAGCCCTGGAATCAAAATCCCTGGGGGCTGATTTGATTATGCCACC 7555
QY 720 AGATTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACT 779
Db 7556 AGATTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACT 7615
QY 780 TCCAGGATTCGTCGCTGCGGATATGTCATATGCTCTTCTTATTCGTATGGCT 839
Db 7616 TCCAGGATTCGTCGCTGCGGATATGTCATATGCTCTTCTTATTCGTATGGCT 7875
QY 840 TGGCAGCTGTCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGTGTGTCAG 899
Db 7676 TGGCAGCTGTCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGTGTGTCAG 7735
QY 900 CAACGTTTTGAACGCTGAACAGCTGCGCGGATTAACCTTAATCTGGATTAAGCCCGTCGA 959
Db 7736 CAACGTTTTGAACGCTGAACAGCTGCGCGGATTAACCTTAATCTGGATTAAGCCCGTCGA 7795
QY 960 TTGCTGTGCGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGTC 1019
Db 7796 TTGCTGTGCGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGTC 7855
QY 1020 CTGACTGTGTAATTTGTCCTTTGAATCCATGCGGATTTAAACGCTGGGATTTGCTGCGGA 1079
Db 7856 CTGACTGTGTAATTTGTCCTTTGAATCCATGCGGATTTAAACGCTGGGATTTGCTGCGGA 7915
QY 1080 AGAGAAGATGACTGGGAGTCCGACCTGTATGCTGGGGCTGACACAGCCTCTTGCA 1139
Db 7916 AGAGAAGATGACTGGGAGTCCGACCTGTATGCTGGGGCTGACACAGCCTCTTGCA 7975
QY 1140 GATAACCGGATATAACCGGAAACTTTCAGAACTCTTTCGCGCCACGACGATGGGACT 1199
Db 7976 GATAACCGGATATAACCGGAAACTTTCAGAACTCTTTCGCGCCACGACGATGGGACT 8035
QY 1200 ATTTATGTCATCTGAAAGCCCGGTTGGAACACAGATCCTCTGGCTTCCGCGAAGAT 1259
Db 8036 ATTTATGTCATCTGAAAGCCCGGTTGGAACACAGATCCTCTGGCTTCCGCGAAGAT 8095
QY 1260 ATCAGGGAAGCTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
Db 8096 ATCAGGGAAGCTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 8155
QY 1320 GGAGGCAATATTTGTTGAAAGCAGATGTCAGCCTCTCTGAAATATGATTTGGCGCA 1379
Db 8156 GGAGGCAATATTTGTTGAAAGCAGATGTCAGCCTCTCTGAAATATGATTTGGCGCA 8215
QY 1380 GGGCCTGATGTCACCTGTGGAGAGAGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1439
Db 8216 GGGCCTGATGTCACCTGTGGAGAGAGAGGATGAGGATGAGGATGAGGATGAGGATGAG 8275
QY 1440 GGAACCGCAATATACCATCAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAG 1489
Db 8276 GGAACCGCAATATACCATCAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAG 8325

AB011549 92721 bp DNA circular BCT 27-APR-1999
LOCUS Escherichia coli plasmid p0157 DNA, complete sequence.
DEFINITION
AB011549
ACCESSION
AB011549.2 GI:4589740
VERSION
ToxR-regulated lipoprotein; tagA.
KEYWORDS
Escherichia coli (strain:O157:H7, sub_strain:RMD 0509952)
SOURCE
Escherichia coli
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
REFERENCE
1 (sites)
AUTHORS
Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K., Yutsudo,H.C., Kubota,Y., Yamauchi,Y., Iida,T., Yamamoto,K., Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S. and Shinagawa,H.
TITLE
Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic Escherichia coli O157:H7 derived from Sakai outbreak
JOURNAL
DNA Res. 5 (1), 1-9 (1998)
MEDLINE
98290540
REFERENCE
2 (bases 1 to 92721)
AUTHORS
Makino,K.
JOURNAL
Submitted (24-FEB-1998) Kozo Makino, Research Institute for Microbial Diseases, Osaka University, Molecular Microbiology; Yamadaoka, 3-1, Suita, Osaka 562, Japan
TITLE
(E-mail:makino@bksn01.biken.osaka-u.ac.jp, Tel:81-6-879-8318, Fax:81-6-879-8320)
COMMENT
On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES
source
1..92721
/organism="Escherichia coli"
/strain="O157:H7"
/sub_strain="RMD 0509952"
/db_xref="taxon:562"
/plasmid="p0157"
/note="RMD 0509952 is a strain of enterohemorrhagic E. coli, EHEC O157:H7"
gene
join(92527..92721,1..2502)
/gene="tagA"
join(92527..92721,1..2502)
/gene="tagA"
CDS
/codon_start=1
/transl_table=11
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4566293"
/translation="MNTKMERWRTPMKLYLSCITLAPLIGVFSATAADNNSAIYFNTSQPNDLQGSIAAEVKEFAQSILPAHPKEDGSDPHLSLRKSLILVPRPVKADKTPVQVEARDDNNKILGTLTLYPPSSLPDTIHLGVEGGIDFTPHNGTKKILNTVAEYNKLSDAGSGSIHSLTNALVEIHTANGRWVRDIYLPQPDLEGRKMRVSVSSAGSYTYFGDKRVLTVGNTLLFKYNGQWFRSGLENNRITYAQHISAEIPARHIVPGNLVIKQGLSGRLNDIKIGAPGELLHTIDIGMLTPDRDFDAKDEAHREYFOTIPVSRMIVNNYAPLHKEVLPTELLTMDPGNGWHSMTQRIGKELVSGIDNANTGLNLTAGLGNHSPYVVAQLAHSRNGYANGIOVHGSGGGIIVTLSTLGNESHVGHSTAGLHYDVPFGKSVHRSANNNNTGWDGKRIFIPNFPSTNEKSLNQCCEPFDGKFGFDAGGSGPFAANRTMTTPNSAIIQRFENKAVFDSRSTGFSKWNADQMEPEYHTIDAEQITASVNELSKMAELMAEYAVVKVHMNNGWNRNRYIPTASADNRGSLTIDHEAGNYSYLFINGDEKVSQGYKSFVSDGQFWRKEDVVDTRARKPEQFGPVTVLVGYDPEGLTSSYIYPAMYAGTGYTSDDSNLSDNDCLQVDTKEGQLREFLANRANTVMKFNHINPTSOPTGATLVCKNKLIDTKSLTPAPEGLTYVNGOALPAKNECCIVSVNSGKRYCLPVGQRSGYSLPDNIVGQEVYVDSKAKAKVLLSDWDNLSYNRIGRFGVNNPADMKKRWANGQYLDPSKPSMRVYIK"
gene
2589..3464
/gene="etpC"
2589..3464
/gene="etpC"
CDS
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31758.1"
/db_xref="GI:3336999"

```
/translation="MLFSLFSGDGLFIKIDIVLKMILTNNRLLCVILLIAGVQLSVI
HPFWLQAASVPLGSRVSAPEATAVTDQTEERFVFTLGRASPLSGEGRQATMPSLIS
DDLSSGDLVRGILYSSVAEHSVAIFAHNNRQFSLSGVEKVPSTADATISAFSDHIV
INQGTSLPLRYDNTKKNAYDNNNLTVGDVITQDNFRVESFVDFIMFSAVTNNIT
LSGRLIPGKHSLSFYNAGLHNDLAVSVNGSELDRQAQIQIMKQLPELKKEITVE
RQGLYDAFIAVGEN"
3675..5432
/gene="etpD"
CDS
3675..5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31759.1"
/db_xref="GI:3337000"
/translation="MLNEEQYTFQFLSVLDVYGFVAVDMHNGILKVVRSKDAKTSAPP
VASDSPGDEVTVPVPSVNAARDLAPLRQLNDNAGAGSVVHPEPSNVLLMTGR
AAVMKRLMEIVRDKVGNRSVATVPLTVASATDVARLVTELTKEVDKAI PAWMTAK
LVADERTNSVLSGEPISQORLIISIKOLDROEDVQGNPKVILYKAKAKDLVEVLTG
ISSIENDSKSPSTALAKGVYIKSHQETNALILGAPDVIRDLNENVISQDIRRPQ
VLVEAIIAQDQDGLNLGIQWVKNHAGVQFTSGLPITTVQTRQNEILSDSDNSA
LSMFGIAAGFTQGNWAMLLTALSTSSKNILATPSIVTLDNMEATFNVGQEPVLSG
SOTSGDNIFENTVERTVGIKLRVKPQINEGDSVLIEIQEVSQVADTAVATVTDLGA
TENTRTVAMLVGNGETVVGGLDKSIRGSESKVPLIGDIPVLGHLFRKSEQTAK
RNLMFIRPTIIRERDGFHSAEAKYQSFNQEQVQSRGKETTALTFLNEBQLRLSPDQD
DYAFRVRKAAIAFAQEM"
5432..6937
/gene="etpE"
CDS
5432..6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"
/translation="MSRVVQVNSERPLLPFSRQRIILLRQEGNRVFCMEDTPA
SALLEVRRVRAEGLPVNTVTVSAEAFKQLVSSYQQRSDSEARQMAEIGNEMDPTVAGE
LRDRELLDANDADAPIRLINAMLTEAIKEKASDIHETIERHLQVRFKIDGVLREIL
RHLNLSALISIRKVMARDLDAERKVPQDGRVLRIGGRVADVVRVSTPLSNHGERIV
LRLDKNSVSLDLAALGMSQONRHIDALIRPHGILVLTGPTSGSKSTLLAALSLL
NPRDKNIMVEPVEYELGDSIQTVNPKVDMTFAIRAILRQDPDVLVGEIRDE
TAQIAVOASLTGLHLVSLTHTNSAAGALSLODMGIPFPLSTSLIAVAQRLVRLTIC
PCRCOPCOYSTELAMDIPPETTLPQAGCOCSPFGYHGRTGHEILLIDIRTA
IVQGEGLAHTRLAGSRITLTKDGRQKVLAGETSWEEVVRVTSRLQEE"
6939..8162
/gene="etpF"
CDS
6939..8162
/gene="etpF"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"
/translation="MALFHYQASDIHGRKRSGILEADSARHAROLLREOALIPVRIDE
KOVHKKHSRSLRPRPGSSAEALATROLATVAASIPLEALDALLRQSEPRQ
RNLLAAVRKVLGHSLLAAMGNFPTFERLYCANVAAGETSGRLDVLVSLRLADTEQ
RQIMRNLQALLPCVLTAVGVIAILLTAVVPKVEQFIMKQTLPLSRVLGMA
AEVSOYWPWLLPALLAAGIAGRMILHQPQRIAFHLLRLPVPVGRISRLNTARYA
RTLSILNASAVPLLAQAHISGVDLSNDWARHOLATAAELVREGVSLHOLEQTSLPFP
MMRHMIASENGSELDLSMLERAADNQDEFSTQMQLALGLFEPFLVVGAGVVLFIVL
AIIQPLLOLNNMNM"
8193..8627
/gene="etpG"
CDS
8193..8627
/gene="etpG"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31762.1"
/db_xref="GI:3337003"
/translation="MRKQHORGFITLIMVIVILGVLASLVVNLGNKDKADROKV
MSDLVALESTLDMYRLDNRYPTTEQGLRALYKSPVQPEPRYRQDGIIRRLPQDPW
GGDYQLNFGQISDIDIFSPGPDGVPENTEDDIGNWTGLGNAQ"
```

```
gene
8624..9178
/gene="etpH"
CDS
8624..9178
/gene="etpH"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337004"
/translation="MSQRGFTILEMMLILLMGTAGLVLMSPDPSAQNHQOORERL
OALDYALDRSQDGLLMGIOVRFDGMEKFKVLQRTAESPTLAEGDILWQGVWQTW
OPRRAMGKGLPDDVLELQYLRLGLQWMSHDDGAEPDILLPLPGGVTFRLLFRQVG
EBAVGLQVDENGLMTLFEVSL"
9193..9540
/gene="etpI"
CDS
9193..9540
/gene="etpI"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31764.1"
/db_xref="GI:3337005"
/translation="MTLLEVVIVLVFALAGMALMOASTQQAAGIGRMEKVLGWL
DNQVQLQLEKTPENGWGEKTIISFAGTEWYLRWGGPDSDVQPRSRLEVEVRYTKEET
SALVSRSSVVR"
9537..10136
/gene="etpJ"
CDS
9537..10136
/gene="etpJ"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31765.1"
/db_xref="GI:3337006"
/translation="MSQORVKGFTLEMLLALAVFAALSIASFQVLOGGIRAHLSRD
KVQLAEALORGISQMERDLTQMLPRHSRGNELLLAAPHLLKSDDGIGIFTRNSWLP
AGMLPEPELQWVGIVRLRQOKLERLSYFHVDPGSGVSPDVVRLMDGVHAPRLRFFVNGD
WOARMDSTGILFOAVEVTLVMDDFAEPLRFLVSKETA"
10133..11110
/gene="etpK"
CDS
10133..11110
/gene="etpK"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31766.1"
/db_xref="GI:3337007"
/translation="MKLRQGVALLVLLLSLMTVIAAIVIAERNRGTFLRTVAQLDQ
LQAKMDGYAETIAKQIORSQESPRKTHLAQNWASERQFETRGDVRGQIVDAQA
CFNLNAINTVVDVLTSPYAAARIFQOLLINLQVELQARQVTAALRDRDDPKVRG
GAEDVYMGMEPPYLAANQMDQVSELRIQDARLYKLLPYVCLVPLTSDLSVNVN
TLIDSQAPLILAAFLTKPDSLPTVTELLQRRPRTGWESVAAFLDAPLKDITDTSAMPV
LAVSNYFLVRLHVRSGEHLFSQOITLMQWREFRFRIQRYGLTMEVP"
11473..12321
/gene="etpL"
Query Match 99.2%; Score 1476.4; DB 1; Length 92721;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 CTGCAGTCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGTGCCGAGGCC 59
Db 76351 CTGCAGTCGGAGATGAACAGCACCACCTGTGTGTACCCCATCAGCGTGTGCCGAGGCC 76410
QY 60 ATGATTTTGTACAGACTCAATGACTACCGGACGACCTGAACCTTCCGTTCTTCCTCC 119
Db 76411 ATGATTTTGTACAGACTCAATGACTACCGGACGACCTTCCGTTCTTCCTCC 76470
QY 120 AGCCAGTTAAGCCAGCGGTTTCCTGCTGTAATAATGTCGGCAAAACGGGAAGCATCAGA 179
Db 76471 AGCCAGTTAAGCCAGCGGTTTCCTGCTGTAATAATGTCGGCAAAACGGGAAGCATCAGA 76530
QY 180 AGGGCGGGGAACTCCGTCGGCCAGTGAACCGTGCACACTCCGGGCGAGTACATCCCGC 239
```



```
|||||
76531 AGGCGGGGAACCTCCGTCGGCCAGTGAACCGTGCACACTCCGGGACATACATGCCG 76590
QY 240 CGGCGCTGATACCGGCAAGATGGTCCCAACTCCCGCTCCGTCGACGGGCTATTTCAG 299
Db 76591 CGGCGCTGATACCGGCAAGATGGTCCCAACTCCCGCTCCGTCGACGGGCTATTTCAG 76650
QY 300 GATACCTTCGTCATCAACAGCTACAAACAGAGACACAGCTTTTGTCTCAGATCCA 359
Db 76651 GATACCTTCGTCATCAACAGCTACAAACAGAGACACAGCTTTTGTCTCAGATCCA 76710
QY 360 CAAGAAGGGAATATTACAGTCTGCGCAGCACTCAACGGCATCTCAGTTGGCGCTTGA 419
Db 76711 CAAGAAGGGAATATTACAGTCTGCGCAGCACTCAACGGCATCTCAGTTGGCGCTTGA 76770
QY 420 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 479
Db 76771 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 76830
QY 480 AATAAGTTGACTTTCTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTCTGATG 539
Db 76831 AATAAGTTGACTTTCTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTCTGATG 76890
QY 540 ATAAAAAATCTTCTCTGTTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCT 599
Db 76891 ATAAAAAATCTTCTCTGTTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCT 76950
QY 600 GTAGCGCTGATAAAAAAGACATCAAAATTTCTACTATCAGAAACACTGGATTAACT 659
Db 76951 GTAGCGCTGATAAAAAAGACATCAAAATTTCTACTATCAGAAACACTGGATTAACT 77010
QY 660 COTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAACC 719
Db 77011 COTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAACC 77070
QY 720 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGCACAACT 779
Db 77071 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGCACAACT 77130
QY 780 TCCAGGATTTGGTCCCTGCGGATATGATGATGATGATGATGATGATGATGATGATGATG 839
Db 77131 TCCAGGATTTGGTCCCTGCGGATATGATGATGATGATGATGATGATGATGATGATGATG 77190
QY 840 TGGCAGCTGCGGCAACATACAGGACATATGATGCGGGGAGCGCCAGTGGTGTGAC 899
Db 77191 TGGCAGCTGCGGCAACATACAGGACATATGATGCGGGGAGCGCCAGTGGTGTGAC 77250
QY 900 CAACGTTTGAACCGCTCAACAGCTGGCCGATACAGTTAATCTGGATAAAGCCCGTCA 959
Db 77251 CAACGTTTGAACCGCTCAACAGCTGGCCGATACAGTTAATCTGGATAAAGCCCGTCA 77310
QY 960 TTGCTGTGCGCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGAGACCTGATGGTC 1019
Db 77311 TTGCTGTGCGCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGAGACCTGATGGTC 77370
QY 1020 CTGACTGGTATATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA 1079
Db 77371 CTGACTGGTATATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA 77430
QY 1080 ACAGAAGATGACTGGGAGTCCGACCTGTATCTAGTGGGGCTGACAAACGCTCTTTGCA 1139
Db 77431 ACAGAAGATGACTGGGAGTCCGACCTGTATCTAGTGGGGCTGACAAACGCTCTTTGCA 77490
QY 1140 GATAACCGGGATATAAAGGGAACCTTTCAGAAACCTCTTGGCGCAGCAGATGGGACTT 1199
Db 77491 GATAACCGGGATATAAAGGGAACCTTTCAGAAACCTCTTGGCGCAGCAGATGGGACTT 77550
QY 1200 ATTTATGATCAATCTGAAGGCCCGGCGGAAACAGATCTCTGCGCTTCGCGCAAGAT 1259
Db 77551 ATTTATGATCAATCTGAAGGCCCGGCGGAAACAGATCTCTGCGCTTCGCGCAAGAT 77610
QY 1260 ATCAGGGAGCTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
|||||
```

```
Db 77611 ATCAGGGAGCTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 77670
QY 1320 GAGAGGCAATACATTTGGTAAAGACATGTTGCAGCTCTCTGAAAAATGATTTGGCGCA 1379
Db 77671 GAGAGGCAATACATTTGGTAAAGACATGTTGCAGCTCTCTGAAAAATGATTTGGCGCA 77730
QY 1380 GGGCTGTATGTTGACCTGTGTGAGGAGCAGGACTGGGATGGAATAAATGTTGGTACA 1439
Db 77731 GGGCTGTATGTTGACCTGTGTGAGGAGCAGGACTGGGATGGAATAAATGTTGGTACA 77790
QY 1440 GGAACGGCAATATATACCATCACAGTGGCTTGGAGGAGCTGGTGCAC 1489
Db 77791 GGAACGGCAATATATACCATCACAGTGGCTTGGAGGAGCTGGTGCAC 77840

RESULT 5
AX191725 92721 bp DNA linear PAT 15-AUG-2001
LOCUS Sequence 7 from Patent WO0149775.
DEFINITION AX191725
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS Escherichia coli.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 92721)
AUTHORS Iversen, P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source
1..92721
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 24870 a 20267 c 23868 g 23716 t
ORIGIN

Query Match 99.2%; Score 1476.4; DB 6; Length 92721;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTGCACTCCGGAGATGAA-AGCACCACCTGTGTGTACCCCATCAGCTGGTCCCGCAGGCC 59
Db 76351 CTGCACTCCGGAGATGAAAGCAGCACCCTGTGTGTACCCCATCAGCTGGTCCCGCAGGCC 76410
QY 60 ATGATTTTGTGCACAGCTCAATGACTACCGGAGCAGCTGAACCTTCCGGTTGTTTCTCC 119
Db 76411 ATGATTTTGTGCACAGCTCAATGACTACCGGAGCAGCTGAACCTTCCGGTTGTTTCTCC 76470
QY 120 AGCCAGTTAAGCAGCGGTTTCCCTGCTGAAAAATGTGCGGAAAAACGGGGAAGCATCAGA 179
Db 76471 AGCCAGTTAAGCAGCGGTTTCCCTGCTGAAAAATGTGCGGAAAAACGGGGAAGCATCAGA 76530
QY 180 AGGCGGGGAACTCCGTCGCGCAGTGAACCTGCAACACTCCGGGAGTACATGCGCC 239
Db 76531 AGGCGGGGAACTCCGTCGCGCAGTGAACCTGCAACACTCCGGGAGTACATGCGCC 76590
QY 240 CGGCGCTGATACCGGCAAGATGGTGCAGAACTCCCGCTCCGTCGAGCGGGCTATTTCAG 299
Db 76591 CGGCGCTGATACCGGCAAGATGGTGCAGAACTCCCGCTCCGTCGAGCGGGCTATTTCAG 76650
QY 300 GATACCTTCGTCATCAACAGCTACAAACAGAGACAGCTTTTGTCTCAGATCCA 359
Db 76651 GATACCTTCGTCATCAACAGCTACAAACAGAGACAGCTTTTGTCTCAGATCCA 76710
QY 360 CAAGAAGGGAATATTACAGTCTGCGCAGCACTCAACGGCATCTCAGTTGGCGCTTGA 419
Db 76711 CAAGAAGGGAATATTACAGTCTGCGCAGCACTCAACGGCATCTCAGTTGGCGCTTGA 76770
QY 420 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 479
Db 76771 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTTGTGCATC 76830
```


QY	480	ATAAAGTTGACTTTTGTATACAAACATGCGAATTTCCCTTAATCCGGAGCTATTCTGTATG	539
Db	76831	ATAAAGTTGACTTTTGTATACAAACATGCGAATTTCCCTTAATCCGGAGCTATTCTGTATG	76890
QY	540	ATAAATAAATCTTCCCTGTTCTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCT	599
Db	76891	ATAAATAAATCTTCCCTGTTCTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCT	76950
QY	600	GTAGCGCTGTATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAAC	659
Db	76951	GTAGCGCTGTATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAAC	77010
QY	660	CTCTGAGATTACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAC	719
Db	77011	CTCTGAGATTACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTATGCCAC	77070
QY	720	AGATTTCACACCTGGATGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAAC	779
Db	77071	AGATTTCACACCTGGATGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAAC	77130
QY	780	TCCAGGATGTTGGTCCCTGCGGATGATGCTATTATGTCCTTTCTTATTCGTATGGCT	839
Db	77131	TCCAGGATGTTGGTCCCTGCGGATGATGCTATTATGTCCTTTCTTATTCGTATGGCT	77190
QY	840	TGCGACGTTGCGGGAACATACAGCATATGATGCGCGGGAGCGCCAGTGGTGTAC	899
Db	77191	TGCGACGTTGCGGGAACATACAGCATATGATGCGCGGGAGCGCCAGTGGTGTAC	77250
QY	900	CAAGCTTTTGAACCGCTGAACAGCTGCGCGGATACGTTAATCTGATAAACCCGTCGA	959
Db	77251	CAAGCTTTTGAACCGCTGAACAGCTGCGCGGATACGTTAATCTGATAAACCCGTCGA	77310
QY	960	TTGCTGTGCGGAGTCAAGAAAAATACGCTCCAGTATTTCTTGGGGAGACCTGATGTC	1019
Db	77311	TTGCTGTGCGGAGTCAAGAAAAATACGCTCCAGTATTTCTTGGGGAGACCTGATGTC	77370
QY	1020	CTGACTGTTAATTTGTCCTTGAATCCATGGGATTTAAACCGCTGGATTTCTGTCGCG	1079
Db	77371	CTGACTGTTAATTTGTCCTTGAATCCATGGGATTTAAACCGCTGGATTTCTGTCGCG	77430
QY	1080	AGAGAAGATGACTGGGAGTCGACCTGGTATATCTGGGGCTTGACAAACGCTCTTGCA	1139
Db	77431	AGAGAAGATGACTGGGAGTCGACCTGGTATATCTGGGGCTTGACAAACGCTCTTGCA	77490
QY	1140	GATAACCGGGATATAAAGCGGAACTTCAGAAACCTCTTGGCGCCACGAGATGGAC	1199
Db	77491	GATAACCGGGATATAAAGCGGAACTTCAGAAACCTCTTGGCGCCACGAGATGGAC	77550
QY	1200	ATTATGTCATTCCTGAAGCCCGGTTGAAACACAGATCCCTCTGCTTCCCGGAAAGAT	1259
Db	77551	ATTATGTCATTCCTGAAGCCCGGTTGAAACACAGATCCCTCTGCTTCCCGGAAAGAT	77610
QY	1260	ATCAGGAAGCTTTTTCAGTATGGCCATGGATGATGAGGAGCTGTGGCCCTGATCGG	1319
Db	77611	ATCAGGAAGCTTTTTCAGTATGGCCATGGATGATGAGGAGCTGTGGCCCTGATCGG	77670
QY	1320	GGAGGCAATACATTTGGTAAAGCAGATGTCGAGGCTCTCTGAAATATGATTGGCGCA	1379
Db	77671	GGAGGCAATACATTTGGTAAAGCAGATGTCGAGGCTCTCTGAAATATGATTGGCGCA	77730
QY	1380	GGGCTGTATGTCACCTGTGGAGGAGCAGGAGCTGGGATGGAATAAATGTTGGTACA	1439
Db	77731	GGGCTGTATGTCACCTGTGGAGGAGCAGGAGCTGGGATGGAATAAATGTTGGTACA	77790
QY	1440	GGAAACGGCAATATACATCACCAGTGGCTTGGAGGAGCCTGGTCGAC	1489
Db	77791	GGAAACGGCAATATACATCACCAGTGGCTTGGAGGAGCCTGGTCGAC	77840

RESULT 6
ECHECCTP
LOCUS

BCEHECCTP 2407 bp DNA linear BCT 12-DEC-1996

DEFINITION	E.coli 0157:H7 katp gene for EHEC-catalase/peroxidase.
ACCESSION	X89017
VERSION	X89017.1 GI:1580760
KEYWORDS	EHEC-catalase/peroxidase; katp gene; Plasmid.
SOURCE	Escherichia coli.
ORGANISM	Escherichia coli.
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
AUTHORS	Brunder, W.
TITLE	Direct Submission
JOURNAL	Submitted (22-JUN-1995) W. Brunder, Universitaet Wuerzburg, Institut fuer Hygiene und Mikrobiologie, Josef-Schneider-Strasse 2, D- 97080 Wuerzburg, FRG
REFERENCE	2 (bases 1 to 2407)
AUTHORS	Brunder, W., Schmidt, H. and Karch, H.
TITLE	Katp, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic Escherichia coli 0157:H7
JOURNAL	Microbiology 142 (Pt 11), 3305-3315 (1996)
MEDLINE	97124214
PUBMED	8969527
FEATURES	Location/Qualifiers
source	1..2407
	/organism="Escherichia coli"
	/strain="EDL933"
	/sub_species="serotype 0157:H7"
	/db_xref="taxon:562"
	/cell_line="plasmid p0157"
RBS	119..122
gene	131..2341
	/gene="katp"
CDS	131..2341
	/gene="katp"
	/codon_start=1
	/transl_table=11
	/product="EHEC-catalase/peroxidase"
	/protein_id="CAA61429.1"
	/db_xref="GI:1580761"
	/db_xref="SPTREMBL:P77038"
	/translation="MIKKTLPVLLIALSGSPSTAVADKKETQNYFPETLDLTPLR LHSPSPNWDADFDYATRFQOLDLEALKDKIDLLTSDWMPADYGHVGPFFIRMA HCAGTYRYDGRGASGCGOOREPLNSPNDVNLKARLLWPVKKYGSSISWGLM VLGNVLESWGFKTLGAGREDDESOLVYWDGNKPLADNRNKGKLPKLAATQ MGLIYNVEGPGGKPLDASKDIREFSRMAMDEETVALIAGHTFGKHAASPE KIGAGPDGAPVEEQGLGWKNKCTGNYTITSGLEAWSTSPQFTMQYIKNLYKY EWELHSPAGAYOWKPKAANTVDAHDPVSLHMFMTDIALKVDYPKYKTTTFL NDPKAFEOAFARAFWEKLTHRDMPAARYLGNVPAESFTWQDPLPADYTMIDKDI SLKEQVMDLGIPIASELKTAWASASTFRVTDYRGNGGARIRLQPEINWEVPEKLK KVLASITSLQREFNKKQSGKVSADLILVLSGNAIEDAARAGVELEIPETPGRTD ASQEQTDVASFVLEPTADGFRNYYSKSKSHISPVESLIDKASQDLQVPTPTALGG LRVMDINTNNSLGVFTDTPGVLDNKKFFVNLDMSTRWKSKADKEDTYNGFDRKTGALK WKASSVDLIFSSNPFLRAVAEYVSDARNKTIHDFVKSWNKVMNSDRFDLNNK"
sig_peptide	131..199
	/gene="katp"
BASE COUNT	691 a 480 c 607 g 629 t
ORIGIN	
Query Match	72.6%; Score 1081.4; DB 1; Length 2407;
Best Local Similarity	99.9%; Pred. No 1e-311;
Matches 1082; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	407 GTTGGCGCTTGAACCCCTTAGTATTTTGTCTGTAGTATCTATCCCAGCAATAGGTAT 466
Db	1 GTTGGCGCTTGAACCCCTTAGTATTTTGTCTGTAGTATCTATCCCAGCAATAGGTAT 60
QY	467 ATCCCTGTCATCAATAAAGTTGACATTTTGTATACAAACATGCGAAATTCCTTAAATCCGG 526
Db	61 ATCCCTGTCATCAATAAAGTTGACATTTTGTATACAAACATGCGAAATTCCTTAAATCCGG 120
QY	527 AGCTATTCGTATGATAAAAAAACTCTCTCTCTCTGATTTCTCTCGCGCTATCGGGAG 586
Db	121 AGCTATTCGTATGATAAAAAAACTCTCTCTCTCTGATTTCTCTCGCGCTATCGGGAG 180

QY	587	CTTTTCTACCGCTGACCGCTGATATAAAAGAGACTCAAAATTTCTACTATATCCAGAAC	646
Db	181	CTTTTCTACCGCTGACCGCTGATATAAAAGAGACTCAAAATTTCTACTATATCCAGAAC	240
QY	647	ACTGGATTAACTCTCTGAGATTACACAGCCCTGRATCAAAATCCCTGGGGGCTGATTT	706
Db	241	ACTGGATTAACTCTCTGAGATTACACAGCCCTGRATCAAAATCCCTGGGGGCTGATTT	300
QY	707	TGATTATGCCACCAAGATTTCACACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGA	766
Db	301	TGATTATGCCACCAAGATTTCACACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGA	360
QY	767	TTTCTGACAACTTCCACGAGATTGGTCCCTCGGATATATGGTCAATATGGTCTTCTT	826
Db	361	TTTCTGACAACTTCCACGAGATTGGTCCCTCGGATATATGGTCAATATGGTCTTCTT	420
QY	827	TATTCGTATGCTTGGCAGCGTCCGGAACATACAGACATATGATGGCGGGAGCGCC	886
Db	421	TATTCGTATGCTTGGCAGCGTCCGGAACATACAGACATATGATGGCGGGAGCGCC	480
QY	887	CAGTGGTGGTCAGCAACGTTTGAACCGCTGAAACAGCTGGCCGATATAGCTTAATCTGA	946
Db	481	CAGTGGTGGTCAGCAACGTTTGAACCGCTGAAACAGCTGGCCGATATAGCTTAATCTGA	540
QY	947	TAAAGCCGCTGATTCGTGGCAGTCACAGAAAAATACGGCTCCAGTATTTCTTGGGG	1006
Db	541	TAAAGCCGCTGATTCGTGGCAGTCACAGAAAAATACGGCTCCAGTATTTCTTGGGG	600
QY	1007	AGACCTGATGCTCTGACTGATGATGTCGCCCTGATCATCGGATTTAAACGCTGGG	1066
Db	601	AGACCTGATGCTCTGACTGATGATGTCGCCCTGATCATCGGATTTAAACGCTGGG	560
QY	1067	ATTTCGTGGCGGAAGAGAGATGACTGGGAGTCGGACCTGGTATATCGGGGCTGACAA	1126
Db	661	ATTTCGTGGCGGAAGAGAGATGACTGGGAGTCGGACCTGGTATATCGGGGCTGACAA	720
QY	1127	CAAGCCCTTTCAGATAACCGGATATAACAGCGGAACCTTCAGAACCTTTCGCCGCCAC	1186
Db	721	CAAGCCCTTTCAGATAACCGGATATAACAGCGGAACCTTCAGAACCTTTCGCCGCCAC	780
QY	1187	GCAGATGGGACTTATTTATGCTCAATCCTGAAGCCCGCTGGAACACAGATCCTCTGCC	1246
Db	781	GCAGATGGGACTTATTTATGCTCAATCCTGAAGCCCGCTGGAACACAGATCCTCTGCC	840
QY	1247	TTCCGCGAAAGATATCAGGAACCTTTTACATGATGCGCATGATGATGAGGAGACTGT	1306
Db	841	TTCCGCGAAAGATATCAGGAACCTTTTACATGATGCGCATGATGATGAGGAGACTGT	900
QY	1307	GGCCCTGATCGCGGAGGCGATACATTTGTGAAGCACATGTCAGCGCTCTCCTCAAAA	1366
Db	901	GGCCCTGATCGCGGAGGCGATACATTTGTGAAGCACATGTCAGCGCTCTCCTCAAAA	960
QY	1367	ATGATTTGGCGCAGGCGCTGATGGTCACTGTGGAGGAGCAGGACTGGGATGGAAGAA	1426
Db	961	ATGATTTGGCGCAGGCGCTGATGGTCACTGTGGAGGAGCAGGACTGGGATGGAAGAA	1020
QY	1427	TAAATGTTGACAGGAACGGCAATATACCATCACCAGTGGCTGGGAAGGAGCCTGGTC	1486
Db	1021	TAAATGTTGACAGGAACGGCAATATACCATCACCAGTGGCTGGGAAGGAGCCTGGTC	1080
QY	1487	GAC 1489	
Db	1081	GAC 1083	

RESULT 7

AE013690/c
LOCUS 12268 bp DNA linear BCT 26-JUL-2002
DEFINITION *Yersinia pestis* KIM section 90 of 415 of the complete genome.
ACCESSION AE013690 AE009952
VERSION AE013690.1 GI:21957603
KEYWORDS
SOURCE *Yersinia pestis* KIM.

ORGANISM

Yersinia pestis KIM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

REFERENCE
AUTHORS

1 (bases 1 to 12268)
Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Stratley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.

TITLE
JOURNAL

Genome Sequence of *Yersinia pestis* KIM
J. Bacteriol. 184 (16), 4601-4611 (2002)

PUBMED
REFERENCE
AUTHORS

12142430
2 (bases 1 to 12268)
Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Stratley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.

TITLE
JOURNAL

Direct Submission
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA

FEATURES
source

1..12268
Location/Qualifiers
/organism="Yersinia pestis KIM"
/strain="KIM"
/db_xref="taxon:187410"
complement(110..2365)
/gene="katG"
/note="y0870"
complement(110..2365)
/gene="katG"
/function="enzyme: detoxification"
/note="Residues 37 to 748 of 751 are 55.68 pct identical
to residues 26 to 725 of 726 from E. coli K12: B3942;
residues 15 to 751 of 751 are 100.00 pct identical to
residues 1 to 737 of 737 from GenPept:
>gI|AAD37313.1|AF135170_4 (AF135170) catalase/peroxidase
[Yersinia pestis]"
/codon_start=1
/transl_table=11
/product="catalase; hydroperoxidase HPI(1)"
/protein_id="AA084455.1"
/db_xref="GI:21957604"

/translation="MLKNSCLLVKGLDYLMLKLPVLITLAIHVHTPTAWAEPKPTD
SFYLPKSLDLSPLRHNTESNPYQDFNAQOFTLDEAVKPKDKITVLTTSDQWPA
DYNGYPPFFIRMAWHGAGTYRIYDGRGAGGQORFEPPLNSPDNANLDRLLWPI
KKYKAKISWGLDMLVLTGNVALESFGFKTLGFAGREDQSDLYWGAAGKMLSDNR
DKNGKLPPLAATOMGLIYVNPFGPKDPVAAAKDIREFAPMAHDEETVALLAG
CHTGTGAHGAASPEKCLGAPGAGLQOGLGWANKCGNGDKITITSGLEGATTTDP
THFTQYLSNLYKHFWLTKSPAGAWQPKNAANVYPDATPTKPHLMFTTDLAL
KVDPYKKITTRFLENPEEFKMAFAWFKLTHRDMPAARYLGDEVPKFTFWQDPL
PAANYKIDSADISLKDILKTLGSDTKLITKAWASATFRGTDPRGGDNGARILA
PKQWPNVNDPAELHSLVAALMEVONNFNDRSGKVSLSDLIVLGNAALIEAAKA
GYSISIPFTPGTDAEOETDVSSFAVLEPTADGFRNYDAKNTLSPIASLIDRANK
LELYPWTVLIGGLRLVDVNSGSKAGVLTNTPGQUNNFFVNLDMSTKWTCKSPA
EGYFDGDRKTKLTKWTASSVDLVFGSNPELRAVAEVIYASDDAKEFEVHDFTKWEKV
MNLDRFDIKNN"
2811..3800
/gene="y0871"
2811..3800
/gene="y0871"
/function="putative transport"
/note="sugar (D-ribose?) transport; residues 6 to 329 of
329 are 100.00 pct identical to residues 1 to 324 of 324
from GenPept: >gI|AAD37312.1|AF135170_3 (AF135170)
D-ribose binding protein [Yersinia pestis]"
/codon_start=1
/transl_table=11
/product="putative
ABC transporter"
/protein_id="AA084456.1"
/db_xref="GI:21957605"
/translation="MNINLMVNIYSMRKIMKNKTLKSLCTAALISMAGVAGAAASNGL


```

KCLGAPGAGLEOGLGWANKCGSGNGKDTITSGLEGAWTTDPTHTMQYLSNLYKH
ENWLTSPAGAWMPKNAANVVPDADTETKEFPLMEFTDIALKVDPEYKKTITREFL
ENPEFEKAPAFWAKLTHRDGPPAARYLGDVEPKETFTWODPLPAANYKMTDSDIIS
EKDRLIKLTDGLSDTLIKTAWASASTFRGTDRGGDNGARIRLAPQKOWPNDPAELH
SVLAALMEYQNNFNDRSGKSVLSDLIVLGGNAAIEDAAKAGYSISIPPTPGRTD
ASQETDVSSEAVLEPTADGFRNYDYAKRNTLSPASLIDRANKLELVPEWTVLIGG
LRLVDNCGSGKAGVLTWTPGGLNNFFVLLDMSTKTKSPKAGYFGDYDRKTKGL
KWTASVDLVFNSPFLRAVAEVIASDDAKAEFVHDTFKVWEKVMNLDKRFKNN"
5311..5894
/gene="cybc"
CDS
5511..5894
/gene="cybc"
/gene="cybc"
/note="similar to Escherichia coli cytochrome c-562"
/codon_start=1
/transl_table=11
/product="cytochrome c"
/protein_id="AAD37314.1"
/db_xref="GI:5002123"
/translation="MRKPIIAGVPSLLITSCFASLQDDNMILLIANLGIYSSSTDPT
KYTISLEKRNALDAQAIPKLEGAEDSPEDIRYRHFDFLLIEQIDKTKQWAE
GNIQEVKSVGVINIRNTYHSRY"
6042..6572
/gene="cybB"
CDS
6042..6572
/gene="cybB"
/note="similar to Escherichia coli cytochrome b-561"
/codon_start=1
/transl_table=11
/product="cytochrome b"
/protein_id="AAD37315.1"
/db_xref="GI:5002124"
/translation="MHKYSCLQIGIHWLVLLVLIANSIELRGFAPRSYOPWKKMI
HSCGIALILMLTILQIOLRYPTPIVPKPSPIVGLAHVGHWYILFLIALPIGI
ATLYCRGSSWIAFGLPHAEQAFDLADTLKAYHLLANMSYFVIGLHAAALLHY
VKDNTLLEMPKRRG"
complement(6829..7539)
/note="ORF7: hypothetical protein"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAD37316.1"
/db_xref="GI:5002125"
/translation="WIKDNTYQDIVYKDYQDSYQKIGITGRVLGALFYCEPDSDA
ECRDIYQKNTWAEWPGGAEMLPIATLLAAQSOYQSIHAQSAAPLAETLE
EAWRIFIGYALPAPPQSVYLDKMLVFGDSTLKRNMQORVEVTLKQOEPEDH
FGLMMMAWLAEPQVDPVLLADHELLPWSYRYLALLQSDAGHPFYQGLAQLTIL
AHWQHELOVTPAGVELYR"
complement(7685..8545)
/gene="dmsC"
CDS
complement(7685..8545)
/gene="dmsC"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=11
/product="dimethyl sulfoxide reductase subunit C"
/db_xref="GI:5002126"
/translation="MGSGHWEPLMVFVIGQVVGGFVLGLALIFGLSGRQORV
HRSMFVLYMALAFSTLHLSGPMRPNFNSALSNELAGSFFFAVAGFYW
LLAVLDKPAAGKILWVYVWLVGIVFYVNCRVYSINTVPTWDSIYPLGVLITLI
GGPLGLYLLQADYVNGRAWLQPLMISVLILISVASVIMQASPLIYSYVQASAL
IPDYGLSMVWVRLVILVGLGCWICPLIRGRPMTLVMVLGLTLLVITAGELIGRVFYGL
HMTVGMVGG"
complement(8547..9164)
/gene="dmsB"
CDS
complement(8547..9164)
/gene="dmsB"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=11
/product="dimethyl sulfoxide reductase subunit B"
/protein_id="AAD37318.1"

```

```

/db_xref="GI:5002127"
/translation="MTQYGFYIDSRCTGCKTCELACKDFKNLSPDVSFRRIYAYG
GWQDQNGTWHQNYFYAYLSIACNHCSDPACTKVPSPGAKHNRDDGFFVVDNEDICIGC
RYCHNACPYPAPYDAEKGHMTKDCGYERVAVGKPKICVDSCPIRALDMPIDELRE
KYGDIAETAPLPAAHFTLPNIVLKNPANSRPVDTTGHLANPEV"
complement(9176..11647)
/gene="dmsA"
CDS
complement(9176..11647)
/gene="dmsA"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=11
/product="dimethyl sulfoxide reductase subunit A"
/protein_id="AAD37319.1"
/db_xref="GI:5002128"
/translation="MOLSNLMKMETNESRLTGTVTRRKLVTITLGGVLMATGAFSL
PFSRTARVQSALNPTSDANDKVINSGSRCPRLRMHVADGEIKVVEDNTGDD
DREGLHOVRACLGRSMRRRYNPDRLKYPMKRIGARGEGKFRKISWEAFETAGSM
QNIIEYGNELIYNGTIGTIGTTRSWPPGATLLARLMCCGGYLDYSTAQI
AAGLNYTGGWADGNSPDIENSLVLFNPNPAETRMGGVYLYLEQAREKSNAM
IVIDRYTDAAGREDEWIPRPGTDAALASALAYVMTENLVDPFLDQYCVGYDEK
TLPEGAPENSHYKAYILGOCADTKATPOWAAITGIPAAKITLQREIGSVKPAYIS
OCWGPORHSGENTSRJAMLAAILGNVINGNSGAREGSYSLPVRMPTFENPVKT
SISFELWDAIERGPEMTATRDGVQCKDLDPYIKFIMYASNCLYNQHSEINRTHDI
LQDDKCEMIVVDNMTSSAETADILDPCTASEOMDFCLDASSGNMAYVIFADQVI
KPRFEKNYDITSLARHMGVEQKTEGTEQEMWHRHLYQOQSAIPELPSFEDFRA
PLPVAAGFESHVDPLIEYQLQLTGFHYKARTSTYGNVDVLYKASCLQEMINPMDA
ERRGISNGDMVIRFNGRGVQINAKVTPRMGVVALGEGAWYSPDAKRIDRAGSINV
LTQRPSPKAGNPSHTNLVQVAKA"
complement(12089..12859)
/gene="arad"
CDS
complement(12089..12859)
/gene="arad"
/note="similar to Escherichia coli ribulose epimerase"
/codon_start=1

```

```

Query Match      36.0%; Score 536.6; DB 1; Length 13632;
Best Local Similarity 69.7%; Pred. No. 1.5e-148;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;

QY 429 TATTTTTCGTCTAGTATCTATCCAGCAATAGTAGTATCTCTGTTGCATCAATAAGTT 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3115 TATCAATAATAGATCAATCTATCATTAATAGAGACTACGAATGTATTTTCCAATTT 3174

QY 489 GACTTTTGTATACAAATCGAATTTCCCTTAATCCGGAGCTATTCTGATGATAAAAAA 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3175 AATCATGTTAAAAATAGTTGCTTTTAAAGGGGACTTAT--ATATGTTAAAAAAA 3232

QY 549 ACTCTTCTCTGTTCTGATTTCTTGGCGGTATCGGGGAGCTTTTCTACCGCTGTAGCCGT 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3233 ATCTTACCGGTACTAATAACTCTCGCCATTTGTACATAATACACCTACGGCTGGCGAGCA 3292

QY 609 GATAAAAAGACACTCAAAATTTCTACTATCCAGAACACTGGATTAATCTCTCTGAGA 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3293 GAAGCACCACAGACTGATAGTTTATTACCCAAAAGCGCTGCTCTCCACCTCGT 3352

QY 669 TTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGTATTATGCCACACAGATTCAA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3353 TTACATAATAGAGTCCATCTCTACGGAAGAGGATTTTAAATATCTCAACAGTTTAA 3412

QY 729 CAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCCAAGAT 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3413 ACACCTGGATCTTGAAGCAGTAAAGAAAGATATAAAACAGTCTTACCACATCACAGAC 3472

QY 789 TGGTCCCTCGCGGATATGTCATTTGCTTCTTCTTTTATTCGTATGCTGGCAGCT 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3473 TGGTGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3532

QY 849 GCGGGAACATACAGGACATATGATGCGCGGGGAGCGCAGTGGTGGTGGTGGTGGTGGTGGT 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3533 GCGGGNACTACCGCATATATGATGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3592

```

QY	909	GAACCGCTGAACAGCTGCGCGGATACGTTAACTGTGATAAAGCCGCTCGATGCTCTGG	968
Db	3593	GAAGCACTCATAGCTGCCAGATACCGCAACTGTATTAACCGGCTCGGCTTCTGTGG	3652
QY	969	CCAGTCAAGAAAAAATACGCTCCAGTATTTCTGGGAGACCTGATGTCCTGACTGTT	1028
Db	3653	CCTATAAAAAAATATGCGCTTAAATATCTGTTGGCGCCACCTGATGTCCTTACAGGC	3712
QY	1029	AATGTTGCCCTTGATCATCGTGGGATTTAAACGCTGGGATTTCTGCGCGAAGAGAT	1088
Db	3713	AATGCGCGCTCGAATCATGCGGTTTAAACGCTGGGTTTTCAGAGGCGCGTGAGGAT	3772
QY	1089	GACTGGGAGCTCGGACCTGCTATACCTGGGGCTTGACAAAGCCTCTTGCAGATAACCGG	1148
Db	3773	GATTTGGCAATCCGATCTGCTACTGTTGGGCGAGCAACAAGATGCTGTCTGATAACAGG	3832
QY	1149	GATTAACACGGAACTTTCAGAACCTTCTGCGCCACGAGATGGACATTTATGTC	1208
Db	3833	GATAAAATGCAAGCTACCCAAACCGCTGCGCGCGACAAATGGGGTTTGATTTACGTA	3892
QY	1209	RATCTGAAGCCCGCTGGAACACAGATCCTCTGCTTCCGCGAAGATATCAGGGAA	1268
Db	3893	RAACCTGAAGGACCAATATGTTAAACAGACCCGCTTCTGCGGCAAAAGATATCTGTGAA	3952
QY	1269	GCTTTTTCACGCTATGCGCATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGAT	1328
Db	3953	CGGTTTTCAGGATGGCAATGAATGATGAGGAGACCGTGGCATTTGATTGGGAGGCGAT	4012
QY	1329	ACATTTGTTAAGACATGTTGACGCTCTCTGAAATGATTTGCGCGAGGCGCTGAT	1388
Db	4013	ACCTTTCGTTAAGCGCATGCGCGACGCTCTCTGAAATGATTTGCGCGAGGCGCTGAT	4072
QY	1389	GCTGACCTGTGGAGGACGAGGACTGGGATGGGAAATAATGTTGACAGAAACGGC	1448
Db	4073	GAAGCGGGCTAGAACACACAGGTTTAGGATGGCGAATTAATGTTGAGGTAATGTT	4132
QY	1449	AAATATACATCACCAGTGGCTTGAAGAGCCTGGTGCAC	1489
Db	4133	AAGGACGATACCAAGTGGGTTGGAAGGGCATGGACAA	4173
RESULT 9			
LOCUS	AJ414156	220050 bp	DNA linear BCT 06-JUN-2002
DEFINITION	Yersinia pestis strain Co92 complete genome; segment 16/20.		
ACCESSION	AJ414156	AL590842	
VERSION	AJ414156.1	GI:15981150	
KEYWORDS	Yersinia pestis.		
SOURCE	Yersinia pestis.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.		
REFERENCE	1 (bases 1 to 220050)		
AUTHORS	Parkhill,J., Wren,B.W., Thomson,N.R., Titball,R.W., Holden,M.T.G., Prentice,M.B., Sebahia,M., James,K.D., Churcher,C., Mungall,K.L., Baker,S., Basham,D., Bentley,S.D., Brooks,K., Cerdeno-Tarraga,A.M., Chillingworth,T., Cronin,A., Davies,R.M., Davis,P., Dougan,G., Feitwell,T., Hamlin,N., Holroyd,S., Jagels,K., Leather,S., Kariyeh,A.V., Moule,S., Oyston,P.C.F., Quail,M., Rutherford,K., Simmonds,M., Skelton,J., Stevens,K., Whitehead,S. and Barrell,B.G.		
TITLE	Genome sequence of Yersinia pestis, the causative agent of plague		
MEDLINE	Nature 413 (6855), 523-527 (2001)		
PUBMED	21470413		
REFERENCE	1586360		
AUTHORS	2 (bases 1 to 220050)		
TITLE	Direct Submission		
JOURNAL	Parkhill,J.		
COMMENT	Submitted (04-OCT-2001) Submitted on behalf of the Yersinia sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
Notes:	Details of Y. pestis sequencing at the Sanger Centre are available on the World Wide Web.		

FEATURES	(URL, http://www.sanger.ac.uk/projects/Y_pestis/).
Source	Location/Qualifiers
	1..220050
	/organism="Yersinia pestis"
	/strain="CO92"
	/db_xref="taxon:632"
gene	/note="biovar: Orientalis"
	complement(191..715)
	/gene="YPO3215"
	/note="arol"
CDS	complement(191..715)
	/gene="YPO3215"
	/EC_number="2.7.1.71"
	/note="Similar to Escherichia coli shikimate kinase II AROL SW:AROL_ECOLI (P08329) (174 aa) fasta scores: E(): 1.5e-31, 55.4% id in 168 aa, and to Erwinia chrysanthemi shikimate kinase Arol or AROM SW:AROL_ERWCH (P10880) (173 aa) fasta scores: E(): 0, 63.7% id in 168 aa"
	/codon_start=1
	/transl_table=11
	/product="shikimate kinase II"
	/protein_id="CAC92450.1"
	/db_xref="GI:15981151"
	/db_xref="SPTREMBL:O8ZC15"
	/translation="MTQTFMVVGARGAGKTTIGKALQALGVRFVDTDFMQQTSQMT VAEVSEGWDFRLRESMAQAVTAPKTVATGGAVLSSENAFMRDHGRVYLRA SAAVLAKRLAEDPEEAQRPSLTGKPIVEIILDVLAASREALQDVAAHVLDGTQPSLV VEQIQLMLTGENVK"
misc_feature	complement(209..709)
	/gene="YPO3215"
	/note="Pfam match to entry PF01202 SKI, Shikimate kinase, score 245.00, E-value 1.1e-69"
misc_feature	complement(470..544)
	/gene="YPO3215"
	/note="PS01128 shikimate kinase signature."
misc_feature	complement(668..691)
	/gene="YPO3215"
	/note="PS00017 ATP/GTP-binding site motif A (p-loop)."
gene	complement(962..1519)
	/gene="YPO3216"
CDS	complement(962..1519)
	/gene="YPO3216"
	/note="Similar to Rhizobium meliloti phospholipid N-methyltransferase PmtA TR:AAG10237 (EMBL:AF201699) (200 aa) fasta scores: E(): 1.1e-07, 25.9% id in 185 aa, and to Bradyrhizobium japonicum PmtA protein TR:O9LCT2 (EMBL:Y09633) (199 aa) fasta scores: E(): 0.0057, 24.2% id in 178 aa, and to Klebsiella pneumoniae hypothetical 22.7 kDa protein TR:O87755 (EMBL:AJ011907) (196 aa) fasta scores: E(): 0, 52.9% id in 191 aa"
	/codon_start=1
	/transl_table=11
	/product="putative methyltransferase"
	/protein_id="CAC92451.1"
	/db_xref="GI:15981152"
	/db_xref="SPTREMBL:O8ZC14"
	/translation="MLNIVRRKQSFYSYLFQRFASPTVGSVSPSPWLCQAMLNQI DWTQNLNIAELAGDGLTKRLSHMSANSLSQAYEIQPHFVHALQKDSRLQADR SAEQLQEDYDVFCCPLLSIPTKISIRILQOQRRLRANGVLVLFQYSHLSESLLS RYFNWKKIRVVRNFPALYVICHPN"
gene	1787..2002
	/gene="YPO3217"
CDS	1787..2002
	/gene="YPO3217"
	/note="no significant database hits"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="CAC92452.1"
	/db_xref="GI:15981153"
	/db_xref="SPTREMBL:O8ZC13"
	/translation="MIGTDVGIIVVAVNYIAETNDRGRSISCHFFIQSSPNAHPLNVD IKNTASPHVLSIVACSIKKAANKRLS"

```

gene
complement(2249..2770)
/gene="YPO3218"
CDS
complement(2249..2770)
/gene="YPO3218"
/notes="Similar to Escherichia coli hypothetical protein
yafE SW:YAF_ECOLI (P30866) (207 aa) fasta scores: E(): 0,
65.1% id in 169 aa, and to Pseudomonas aeruginosa yafE or
PA3119 TR:087011 (EMBL:U93274) (187 aa) fasta scores: E():
0, 55.2% id in 172 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92453.1"
/db_xref="GI:15981154"
/db_xref="SPTREMBL:O82C12"
/translacion="MQAAADKKLTINIEIQGLAESLPFAQDSFIVISRYSAHHWHV
GKALREVRLPGGVIFMDVVPSPGHVLDIYLVQVEVLRDTSVRNYSFGWLSLF
TEGLVINEVTSRLYLEFSWVARMRTPAHFATAREFQKLASDGVINHAIQAQDS
FTSDIMMIVARD"
complement(3016..3285)
/gene="YPO3219"
CDS
complement(3016..3285)
/gene="YPO3219"
/notes="no significant database matches"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAC92454.1"
/db_xref="GI:15981155"
/db_xref="SPTREMBL:O82C11"
/translacion="MGSVMGISEESIRRLNEKNAVGHSAKWAIIISAVFIIMLF
YKHELVLAGIFVSETTWNKKQKQKYSKNLQIQIEEDKV"
complement(3088..3138)
/gene="YPO3219"
/notes="one of 2 probable transmembrane helices predicted
for YPO3219 by TMHMM2.0"
complement(3151..3216)
/gene="YPO3219"
/notes="one of 2 probable transmembrane helices predicted
for YPO3219 by TMHMM2.0"
complement(3468..4226)
/gene="YPO3220"
complement(3468..4226)
/gene="YPO3220"
/notes="Similar to Escherichia coli hypothetical protein
xGR SW:YGR_ECOLI (P76010) (244 aa) fasta scores: E(): 0,
40.7% id in 241 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92455.1"
/db_xref="GI:15981156"
/db_xref="SPTREMBL:O82C10"
/translacion="MGVSETSKQVFKKNKLAICAILRDLOKNDTAVMVTHARGOFI
SRLLDQPTNDFIFDFGSVENENVLALGAEOLTIIVPTGAKIEFTCKLKHVELS
LPAFSAIPQLFYIQRREYFRVSIQWPAAYCSGFPGTQYKYLADISLGMGLY
AMKSEFPQGSGLVDAADVDCGFLFKLDLOFIRALDKQVYNNKGEMLTVQRLSFK
FRLSPQKGLQRAIFELEKQOTAKARFQEGL"
complement(4554..5813)
/gene="YPO3221"
/notes="proA"
complement(4554..5813)
/gene="YPO3221"
/EC_number="1.2.1.41"
/notes="Similar to Serratia marcescens gamma-glutamyl
phosphate reductase ProA SW:PROA_SERMA (P17857) (417 aa)
fasta scores: E(): 0, 85.1% id in 417 aa, and to
Escherichia coli gamma-glutamyl phosphate reductase ProA
SW:PROA_ECOLI (P07004) (417 aa) fasta scores: E(): 0,
72.4% id in 417 aa"
/codon_start=1
/transl_table=11
/product="gamma-glutamyl phosphate reductase"

```

```

/protein_id="CAC92456.1"
/db_xref="GI:15981157"
/db_xref="SPTREMBL:O82C09"
/translacion="MNLLEHMGKAQAQSWQLAMSLTAKKNQALAVIANLLESOTI
LQANEDMAAARESGMSEALLDRLLTPARLAATANDVQVRLNDPVGRIVDGSLID
SGLKERRVPVIGVIGVYEARNVTVIDVASLCKTGNVILRGGEETHYNTQATVNV
IQRLQCGGLPAAAVQARIESPQRLVNEELRLDRKYDMLIPRGASLHKLCREOSTIP
VITGGIGVCHTFVDNADFEKALLVIENAKIQRPSACNSLETLLHQAVAKTEPLLS
ARHAFGVTLHASPRLAMPYADKRVAVEAADYDDEWLSLDLNDYIVTDIDAAIDH
IREHGTSHSDAILTRLSHAIEYFVRAVDSSAVYVNASTRFTDGGQFGIGAEVAVSTOK
LHARGPMGLDALTYKWIQVGGDDLVR"
complement(4572..5744)
/gene="YPO3221"
/misc_feature
/notes="Pfam match to entry PF00171 aldedh, Aldehyde
dehydrogenase family, score -217.90, E-value 0.00013"
complement(4776..4841)
/gene="YPO3221"
/misc_feature
/notes="PS01223 Gamma-glutamyl phosphate reductase
signature"
complement(5823..6926)
/gene="YPO3222"
CDS
complement(5823..6926)
/gene="YPO3222"

Query Match      36.0%; Score 536.6; DB 1; Length 220050;
Best Local Similarity 69.7%; Pred. No. 3e-148;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;

QY 429 TATTTTGTCTAGTACTATCCAGCAATAGGTATATCTGTTGTCATCAATAAGTT 488
DB 121630 TATCAATAATAGATAAATCTATATTAATAGGACTAACGAATGATATTTTCCATTT 121689

QY 489 GACTTTGTATACAAATCGGAATTTCCCTTAATCCGGAGCTATTCTGTATGATAAAAAA 548
DB 121690 AATCATGTTAAAAAATAGTTGCCTTTTAGTTAAAGGGGACTTAT--ATATGTTAAAAAAA 121747

QY 549 ACTTCTCTGTTCTGATTTCTTCTGGCGCTATCCGGGAGCTTTTCTACCGCTGTAGCGGCT 608
DB 121748 ATCTTACCGCTACTAATACTCTCGCATTTGTACATAATACACTACGCTGCGGACGA 121807

QY 609 GATAAAAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTTAACTCTCTCTGAGA 668
DB 121808 GAAGCACCAAGACTGATAGTTTCTTACTTACCCAAAAGCGCTCTCTCTCCACTCCGT 121867

QY 669 TTACACAGCCCTCAATCAATCCCTGGGGGCTGATTTTCTATGATGACACAGATTTCAA 728
DB 121868 TTACATAATATAGAGTCCATCTCTACGGAAGAGATTTTAATTTATGCTCAACAGTTTAA 121927

QY 729 CAGCTGTATAGGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGAT 788
DB 121928 ACACCTGGATCTTGAAGCAGTAAAGAGATATAAAAAACAGTTCTTACCACATCACAAGAC 121987

QY 789 TGTGTCCTCGCGATTTATGTCATTTATGGTCCCTTCTTTTATTCGTATGGCTTGGCAGGTT 848
DB 121988 TGTGTCCTCTGCTGATTTATGTTATGTTGTCATTTCTTTATTCGTATGGCTTGGCAGGTT 122047

QY 849 GCGGGAACATACAGGACATATGATGCGCGGGAGGCGCCAGTGTGCTGACGCAACGTTTT 908
DB 122048 GCGGGAACCTACCGCATATATGATGCGCGTGGTGGTCCGATGGCGGACGCAAGATTT 122107

QY 909 GAACCGCTGAACAGCTGGCGGATAAAGTTAATCTGGATAAAGCCGCTCGATTCGTCTGG 968
DB 122108 GAGCCACTCAATAGCTGGCCAGATAACGCCAACCTTGATAAAGCGCGCTCGGCTCTCTGG 122167

QY 969 CCAGTCAAGAAAAAATACGCTCCAGTATTTCTTCTGGGACACCTGATGGCTGCTGACTGGT 1028
DB 122168 CCTTAAAAAAGAAATATGCGCTAAAAATATCGTGGGCGACCTGATGTTCTTACAGCG 122227

QY 1029 AATGTTGCCCTTCAATCCATGGGATTTAAAAACGCTGGGATTTGCTGCGGCAAGAGAT 1088
DB 122228 AATGTCGCGCTCGAATCTATGTTGAGGTTTAAAAACGCTGGGTTTTCGAGGAGCGCTGAGGAT 122287

```


[illegible]

```

RESULT 10
AB017595
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB017595
2587 bp
DNA
linear
BCT 06-AUG-1999
Legionella pneumophila gene for catalase-peroxidase, complete cds.
AB017595
GI:4996127
kDa; catalase-peroxidase.
Legionella pneumophila (strain:AM511) DNA.
Legionella pneumophila
Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
Legionellaceae; Legionella.
1 (sites)
Anemura-Maekawa,J., Mishima-Abe,S., Kura,F., Takahashi,T. and
Watanabe,H.
Identification of a novel periplasmic catalase-peroxidase Kda of
Legionella pneumophila
FEMS Microbiol. Lett. 176 (2), 339-344 (1999)
93356723
2 (bases 1 to 2587)
Anemura-Maekawa,J. and Watanabe,H.
Direct Submission
Submitted (10-SEP-1998)
Junko Anemura-Maekawa, National Institute.
of Infectious Diseases, Department of Bacteriology; Toyama 1-23-1,
Sinjuku-Ku, Tokyo 162-8640, Japan (E-mail:jmaekawa@nih.go.jp,
Tel:81-3-5285-1111, Fax:81-3-5285-1163)
location/Qualifiers
1. .2587
/organism="Legionella pneumophila"
/strain="AM511"
/db_xref="taxon:446"
203. .2452
/gene="kda"
203. .2452
/gene="kda"
/codon_start=1
/transl_table=11
/product="catalase-peroxidase"
/protein_id="BAA78342.1"
/db_xref="GI:4996128"
/translation="MEKRRIPPLFAAFTLAISPIPNYAHQDKPKTKYWPVKMLD
LSPURQPNATSNPMGEKFNVAEFSNLDNAVIEDLKIMTSDQWPDYKNGYGLPF
IRMSWHAQAQYTRIDYRGKGANGFQRFAPQSWPNANLDKARLLWPKQYKGRKS
WADLLVLAGNVAMSGMGFKTIGFAGRGDEAWAIINWGPCKWLESKRQDKGRLEK
PLAATVNLGLIYVNPGEKGVDPDLAAAEIRTFTRGMNDEETVALIAGHAFGRKLE

```

BASE COUNT	784 a	551 c	611 g	641 t
ORIGIN				
Query Match	25.8%; Score 384.8; DB 1; Length 2587;			
Best Local Similarity	65.8%; Pred. No. 2.6e-103;			
Matches 592; Conservative	0; Mismatches 302; Indels 6; Gaps 2;			
QY	590	TTTACCGCTGTAGCCGTGTATATAAAAGAGACTCAAAATTTCTACTATTCAGAAAACT	649	
DB	271	TTATGCCATGCCAAGAAGCAAGCCGAAAAACCAATCAATACTTGGTGGCCTAAAAATGCT	330	
QY	650	GGATTTAATCCTCTCAGATTACAGACGCTGAATCAATCAATCCCTGGGGGGCTGATTTGA	709	
DB	331	CGATCTTAGCCGCTCGGGCAACCCATCCACTTCGAATCCCATGGGTGAAAAATTCAA	390	
QY	710	TTATGCCACCAGATTTCAACAGCTGGATGAGGCTCTGAAAAAGATATCAAGATTT	769	
DB	391	CTATGCCGAAGATTCATTAATAGCTTAGATTTGAATGCCGTGTAGAGATCTCAAAAAAT	450	
QY	770	GCTGACAACTTCCAGGATTTGGCCCTCGGATTTATGTCATTTATGCTCTTCTTTAT	829	
DB	451	AATGACTTACCTCCGAAGACTTGGGGCTGCTGATTTGTAATGTAATGTAATGTCAT	510	
QY	830	TCGTATGGCTTGGCAGGCTCCGGAACATACAGACATATGATGCCGGGGAGGCCAG	889	
DB	511	ACGCATGTCATGTCGCGCAGAACTTACCGCATCTATGATGCGCGTGGTGGAGCAAA	570	
QY	890	TGTTGTTGTCAGCAACGTTTGAACCGCTGACAGCTGGCCGGATACGTTAATCTCGATAA	949	
DB	571	CGTGGGCTTCGAGGTTTCCGCCCGCAAAACACTGCGCGATATGATGCGCGTGGTGGAGCA	630	
QY	950	AGCCCGCTCGATTTGTTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCTCTGGGAGA	1009	
DB	631	AGCCAGCAGTTTACTGTGGCCATTAAGCAGAAATATGGCGCAGATTTTCATGGCGCTGA	690	
QY	1010	CCTGATGTCCTGACTGGTAACTGTGCCCTTGAATCCATGGGATTTTAAACCGCTGGATT	1069	
DB	691	TTTGTGGTACTTCTGGAATGTTGGATGGGATCTATGGCTTCCAGACCATTTGGATT	750	
QY	1070	TGCTGGCGGAAGAGAAGATGACTGGGAGTCGGACCTGGTATCTATGGGGCCCTGACAA	1129	
DB	751	TGCTGGAGGGCTGAGACGATGGGAGGCCATCAATATCAATTTGGGGCCCTGAAGGCCAA	810	
QY	1130	GCCTCTTTCAGATAA---CCGGGATAAAACGGGAACCTTCAGAAACCTTGGCCGCCAC	1186	
DB	811	ATGGCTGGAAGAGTAAGCGCCAGGATAAAGATGGAATACTTGAATAACCGCTTGTCTCGAC	870	
QY	1187	GCAGATGGACTTATTTATGTCAATCTGAAGGCCCGGTTGGAAAAACCATCTCTCTGGC	1246	
DB	871	CGTAATGGCTTAATCTAATGTCAATCGGAAGGACCAACGGCGTCCAGATCTCTCTGC	930	
QY	1247	TTCCCGGAAGATATCAGGGAAGCTTTTTCACGATATGCCATGGATGATGAGGAGACTGT	1306	
DB	931	TGCGGCCGAAAAAATTCGAGAGACTTTTGGGCGTATGGCCATGAATGACGAAGAACAGT	990	
QY	1307	GGCCCTGATCGGGGAGGGCATACATTTGGTTAAGCACATGGTGCAGCGCTCTCTGAAA	1366	
DB	991	TGCTTTTAATTCGGGCGGGCAGCGATTCGGAATAACACATGGAGCCCGT---CCGGCAA	1047	
QY	1367	ATGTATTTGGCGCAGGGCCTGTGTGTGACCTTGTGGAGGACGAGGACTGGGATGGAAAA	1426	
DB	1048	ATATCTGGGCCAGCCGCAAGCAGCAGGACATAGAGACAAGAGCTTTGTTGGAAAAA	1107	
QY	1427	TAAATGTGTCAGGAAGCGGAATATACCATCACCAGTGGCCTGGAGGAGCGCTGCTG	1486	

Db	1108	CAGTTACGCGCAGCGGCAAGGAAAGACACAGTACCAGTGGTTAGAGGTCCTGGAC	1167
RESULT 11	AF276752	5249 bp DNA linear	BCT 04-DEC-2000
LOCUS	Legionella pneumophila catalase-peroxidase (kata) gene, complete		
DEFINITION	AF276752		
ACCESSION	AF276752		
VERSION	AF276752.1	GI:11528084	
KEYWORDS			
SOURCE	Legionella pneumophila.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group; Legionellaceae; Legionella.		
REFERENCE	1 (bases 1 to 5249)		
AUTHORS	Bandyopadhyay, P. and Steinman, H.M.		
TITLE	Catalase-peroxidases of Legionella pneumophila: cloning of the kata gene and studies of kata function		
JOURNAL	J. Bacteriol. 182 (23), 6679-6686 (2000)		
MEDLINE	20528318		
PUBMED	11073912		
REFERENCE	2 (bases 1 to 5249)		
AUTHORS	Bandyopadhyay, P. and Steinman, H.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-JUN-2000) Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461-1602, USA		
FEATURES	Location/Qualifiers		
source	1..5249		
	/organism="Legionella pneumophila"		
	/strain="JR32"		
	/db_xref="taxon:446"		
misc_difference	2225		
	/note="compared to Legionella pneumophila kata sequence in GenBank Accession Number AB017595"		
	/replace="g"		
misc_difference	2227		
	/note="compared to Legionella pneumophila kata sequence in GenBank Accession Number AB017595"		
	/replace="c"		
gene	2386..4635		
	/gene="kata"		
CDS	2386..4635		
	/note="kata; localized in periplasm"		
	/codon_start=1		
	/transl_table=11		
	/product="catalase-peroxidase"		
	/protein_id="AAG37106.1"		
	/db_xref="GI:11528085"		
	/translation="MFRTKTLFLAFTLAISSIFPNVYAHQEDKPTNINWPKMLD LSPLOPNATSNPMCKEYNAEENSLDNLVIEDLKLMTTSODWPDYGNVPLF IRMSWHAAGTVRIYDRGGANGQRFAPQNSWDNLDKARELLWPKQVGRKIS WADLLVLGANGVMSHGPKTITGFAGGREDWEAENINWPGKWLSEKQDKGLBK PLAAITVMGLIYVNPSPGVPDPLAAAEKIRETFGRMNDEETVALIAGHAFKTH GAASQGLPAPAAAGIEQGFQWKNSSGKGDITISGLEAWTVPTFWSHNTIQ NLFENWYKTSFGPAIDQWPNSSNASSVPDAPDPKSRHAPVMTLDLALKEDPVYS KIAKRFLOPKFEDAFARAFWFKLTHRDMPRSRLVGLSLVPKEAMIQDVPVVDVYL VANDIANLKGILSLTSELVKTAWASATFTGTDMRGGANGAIRLAPQKDWPA NDQELAKVLKLTESIQNNFNNAQDGKKISLADLVLGNGAATEQAQKAGDYIIVP FTPGRMTDQGMNTQKSEVLEPRADGRPNFTDKSNNSPPEMLVERASLLKLISVPE TTVLGVGMRLNANTQNGYVFTDKPTGLNDFFINLLSMSTEWKSKSEGTVEGYE RKTGLKWKATSVLDLIFGANSERLVAEAYATDDAKKEFIQDFINAWKVTADRFDI KAAANINIS"		
BASE COUNT	1771 a 1058 c 1095 g 1325 t		
ORIGIN			
Query Match	25.8%; Score 384.8; DB 1; Length 5249;		
Best Local Similarity	65.8%; Pred. No. 3.1e-103;		
Matches 592; Conservative	0; Mismatches 302; Indels 6; Gaps 2;		
QY	590	TTCTACCGCTGAGCGCTGATATAAAAAAGACACTCAAAATTTCTACTATCCAGAAACACT	649

/db_xref="GI:19914794"
 /translation="MOOSLSHSLFPLCGRAFORDDLMERNTCMILALDVTDRREALKI
 AEDVWEFVAIKVYPLILATGLIGIIRELAEPFVILAEKFAVIDPNTNLCIDQVFEA
 GADAVLAQPTGSDSDACIEIASYKDFVVFVSEMSHPGADFPQVAEIAKMALE
 AGAFGLVAPATPRKVKRIIRIIGDKLTIIISPCVGAQGGKASDVIAAGADWIVORSI
 YKASPEKARQIAEILQELRGEY"
 5307..7799
 /gene="MA0970"
 CDS
 5307..7799
 /gene="MA0970"
 /codon_start=1
 /transl_table=11
 /product="sensory transduction histidine kinase"
 /protein_id="AA04402.1"
 /db_xref="GI:19914795"
 /translation="MOEYSKINSQKGIQKGMNAKVQTPATNOLVLSVDNNGTPI
 IHSNEASEPTLKMGVSGVEKLPSSIVDLQVWISRSNPEKTEVKGKLIYLVFSPIL
 PEQRVYSIGFDSIDOKREKGMVSETRKARDLVEELRSEKRYLFTDIDEGF
 CVTEVLEDENDCPIDRYFLEINNASFERQGIKNVGRMRREIAPHEEHWFQIYNVA
 LTGEVRFENPALGHHFYDVAFRTGDDPHRRVGILFDIAKRKRAEEMLESABEQ
 AFLQLSDALRPLFDIEIQAVATRVILGSHGJARVAYGEVDDDTFIVFERNYVAPG
 VPLVTRFMADPSPILLSALKEGRTIVVPDLASSTELSEASERSYANIGIASLMGVP
 LVKGFRVLAELVHHVTPQWSPAYVALIETEARTAAVERAKAEALKESEERTL
 LSTSPIGVGVSSAESKIIYTPAYDRILGNPSELIGNSTDVWNPERSKERTL
 KKGVDRETEKPKKDKTPVWVSINTSPISFGGRKAAMGTIQDISEKQAEALRES
 ELKFRILNTELEKVERVELEKAYKFLQEIIRKQIHRKRNKQLVLSLLDQ
 AKFRKKNIEDSKILEAFKESQDRVISMALIHEELHKSGEIDTLNFSAYIHELSEN
 FLSYRLNGDISIDMDEEDIFDDMTSVPLMIVNELVNSLKHAFSDRDKGEIRIK
 LHRKNRESIDECCLAFILSDNGIGIPKDLIEIDIESLGLVLTLIDLDGELE
 LKDDGTETFRSITERNQANSPTQQNEF"
 complement(8223..8672)
 /gene="fur"
 /note="MA0971"
 complement(8223..8672)
 /gene="fur"
 /codon_start=1
 /transl_table=11
 /product="ferric uptake regulation protein"
 /protein_id="AA04403.1"
 /db_xref="GI:19914796"
 /translation="MSNLQSDPDLIIKALRGYKATPQRIAGIOFVLNHAHPTAQR
 IYSEVKVPTVSLATYIKTVQILKEVLIQFNLKDOARFDPNPEHAHLVCLQCK
 SINDCTDPMISEVDLRMSNEVDPSAGENWFNIFGICSNCRKKSICD"
 8913..11123
 /gene="MA0972"
 CDS
 8913..11123
 /gene="MA0972"
 /codon_start=1
 /transl_table=11
 /product="catalase/peroxidase"
 /protein_id="AA04404.1"
 /db_xref="GI:19914797"
 /translation="MGNYMTDDKMSVTSKANKQETGRDMSNRWPNHLKLEILHQ
 HSKSNPMDENFYAEKFLDLAAVKDLAALMDSQDWPADFGHYGLPIRMAWH
 SACTYRAGDGRGGGQGFQFAPLNSPDVNLDRALLWPKIKYGRKISWADMI
 LTGNVAMEWPKTFGGGREDYWEFDQYVWGSDETLWGBERTGTDRDLNPLAAY
 QMGLIIVPEGNPDPIAAKDIREFARMAMNDEETVALIAGHAPKRGTHGAPPA
 SHVGPPEASIAEQGLGWKSGFTGKDDTITGGLEVTVNTPTKWSNFFRILFGY
 EWEITSPAGAYQWPKGGAGACTIPDAHDPKSRHAPSMTTDLSLRFPDVPYKISRH
 FYPENQGLASFAWAFKLTIRDMGPRLYGLPEVPAELIWDQPIPAVNHLEIDSK
 IAFKRLILASGISLSQVSTAWVSASTFRGSKRGANGARIRLAPOKDWVNPAPAE
 LKAVLNTLEGISEFNSASGGKSLADLIVLACGAGVEQAANKNGVDVTPFLFGR
 MDALQOTDVVSFALLEIADGPRNLYKQYPLPARELLVKAQLLTAPETVILVG
 GMVLTNFGHTQHTQKPEALTNFVNLDMGTWKVASYDKVDIFPEGCDRKTGE
 VKWTRVDLIFGNSQLRALAEVYGSADAQEKVQDFVAANTKVMNLDRFDLA"
 BASE COUNT 3077 a 2495 c 2699 g 3049 t
 ORIGIN

Query Match 23.5%; Score 349.4; DB 1; Length 11320;
 Best Local Similarity 66.6%; Pred. No. 1.5e-92;
 Matches 548; Conservative 0; Mismatches 266; Indels 9; Gaps 3;
 QY 672 CACAGCCCTGAATCAATCCCTGGGGCGGTGATTTTGATTATGCCACCAGATTCAACAG 731

9045 CATTCCTCCAAGTCCAAACCCGATGGCGGAGGACTTCAATTACGGAAGGACTTCAAGAGC 9104
 732 CTGGGATATGGAGGCTCTCAAAAAGACATATCAAGATTTTGTGACAACCTTCCCGAGATTGG 791
 9105 CTAGATCTGGCGGCTGTGAAAAAGATCTTCGGCGGCTGATGACCGATTCGCAGACGATGG 9164
 792 TGCCCTGCGGATTATGATTCATTATGTCCTCTTTCTTTATTCGTATGCGTTCGACAGGTGCC 851
 9165 TGGCGCGGGATTTTGGCCATTATGGCCCTTTGTTTATTCGCGATGCGACAGCGCA 9224
 852 GGAACATACAGACATATGATGCGCGGGAGGCGCCAGCTGCTGTCAGCAACGTTTGA 911
 9225 GGCACCTTATCGCGCGGTGACGGCGCGCGGTAGAGGCGACGACGATTCGCG 9284
 912 CGGCTGAACAGCTGGCGCGGATAACCTTAATCTGGATAAAGCCGTCGATTCTTGTGGCCA 971
 9285 CGGCTCACAGTGTGGCGGACATGTCACCTGACAAAGCGGCTGCTCTCTGCGCG 9344
 972 GTCAAGAAAAATACGGCTCCAGTATTTCTTGGGAGACCTGATGTCCTGACTGGTAAT 1031
 9345 ATCAAGCAGAAGTATGGCGGAAAAATTTTCATGGCGCGACCTTATGATTCTTTACGGCAAC 9404
 1032 GTTGCCTTGAATCCATCGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAAGATGAC 1091
 9405 GTCCGAATGGAAACTATGGGATTCAAAACGTTTGGTTTTCGGGGCGCGTCCGCGAGGACGTC 9464
 1092 TGGGAGTCCGACCTCT--GTACTCTGGGGCGCTGCACAAAGACCTCTTTCAGATAAACCGG 1148
 9465 TGGGAGCGCGACCGAGACGCTATTTGGGGCTCTGAGGACACGCTGGCTGGGTGATGACGC 9524
 1149 GATA---AAACGGGAAACTTCAGAAACCTCTTGGCGCCACGACGATGGGACTTATTTAT 1205
 9525 TACACCGCGCAGCGGACCTCGAAAATCCGCTCGCTGCGGTGAGATGGGCTGATCTAT 9584
 1206 GTCAATCTGAAGCCCGCGGTGGAACACACAGATCTCTGCTGCTCGCGGAAAGATATCAGG 1265
 9585 GTCAATCTGGAAGCGCCGACGACGATCCGATCGATAGCGCGCGCCAAAGATATCCGC 9644
 1266 GAAGCTTTTTCAGTATGCCATGGATGATGAGGAGACTGTGGCCCTGATCCCGGAGGG 1325
 9645 GAGGTTTTTCGCTCGCATGCCATGAACACGACGAGACCGTTGCGCTCATCGCTGGCGGT 9704
 1326 CATACATTTGTAAGACACATGGTCACGCTCTCTGAAAATGATTTATGGCGCAGGGCCT 1385
 9705 CACGCTTTCGAAAACCCACGCTGCGG---GCCCTGCGCTCCCATGTGCGGCCCTGACGCA 9761
 1386 GATGTGTCACCTGTGGAGGACGAGGACTGGGATGGAAAAATAATGTTGTGTACAGGAAC 1445
 9762 GAAGCGCGCATATGACGCGCGCAGGTTCTCGTTGGAGAGACGACGCTTCGGCATTGGCNA 9821
 1446 GCAAAATATACCATCACCAGTGGCTTGAAGGAGCGCTGGTCTGA 1488
 9822 GSGCATGACAGCATCACCGCGCGCTGGAGGTGACCTGACCTGGACCA 9864

RESULT 15
 AF294823
 LOCUS
 DEFINITION
 Shigella sonnei O antigen gene cluster, complete sequence;
 insertion sequences IS1 InsB (insB), InsA (insA), IS630
 transposase, IS629 ORF4 (orf4), IS91 transposase (insB) and InsA
 (insA), and IS911 InsB (insB) genes, complete cds.
 AF294823
 ACCESSION
 VERSION
 AF294823.1 GI:15149169
 KEYWORDS
 SOURCE
 ORGANISM
 Shigella sonnei.
 Shigella sonnei
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Shigella.
 1 (bases 1 to 17986)
 Xu D.Q., Cisar J.O., Ambulos N. Jr., Burr D.H. and Kopecko D.J.
 Molecular Cloning and Characterization of Genes for Shigella sonnei

Form I O Polysaccharide: Proposed Biosynthetic Pathway and Stable
Expression in a Live Salmonella Vaccine Vector
Infect. Immun. 70 (8), 4414-4423 (2002)

JOURNAL
PUBMEDREFERENCE
2 (bases 1 to 17986)AUTHORS
Xu, D.O., Cisar, J.O., Ambulos, N. Jr., Burr, D. and Kopecko, D.J.TITLE
Direct SubmissionJOURNAL
Submitted (09-AUG-2000) Oral Infection and Immunity Branch,
National Institutes of Health, National Institute of Dental and
Craniofacial Research, Bldg. 30, Room 302, Bethesda, MD 20892, USAFEATURES
Location/Qualifiers

source

1. .17986
/organism="Shigella sonnei"
/strain="536"
/db_xref="taxon:624"
/clone="cosmid pwr101"
1. .768
/note="similar to Escherichia coli IS1"
/insertion_seq="IS1"
complement(16. .519)
/gene="insB"
complement(16. .519)
/gene="insB"
/note="similar to Escherichia coli IS1 InsB"
/codon_start=1
/transl_table=11
/product="insB"
/protein_id="AAK85161.1"
/db_xref="GI:15149170"

repeat_region

gene

CDS

/translation="WPGNCTHYGRWPHDFTSLKLRPQSVTRIOPGSDVIVCAEMD
EQMGVGAQRWLFAYDSLRKTVAHVGERPMAVGLRLSLSPDVIWNTD
WFLYSRLKGLHVISKRYQRIERHNLRLQHLARLGRKLSLSFSLVLDKDKVIGHY
LNIRHYQ"
complement(438. .713)
/gene="insA"
complement(438. .713)
/gene="insA"
/note="similar to Escherichia coli IS1 InsA"
/codon_start=1
/transl_table=11
/product="insA"
/protein_id="AAK85162.1"
/db_xref="GI:15149171"

gene

CDS

/translation="WASVISCPSCSATDGVVRNGKSTAGHORYLCSHCRTWQLQFT
YASQGRHQKIIDMAMNGVGCATARIMGVLNLRHLKNGRSR"
complement(721. .724)
complement(734. .739)
complement(756. .761)
788. .1720
/gene="wzz"
788. .1720
/gene="wzz"
/note="similar to Shigella sonnei rfbR and Actinobacillus
actinomycetemcomitans wzz proteins"
/codon_start=1
/transl_table=11
/product="truncated Wzz protein"
/protein_id="AAK85163.1"
/db_xref="GI:15149172"

RBS

-10_signal

-35_signal

gene

CDS

/translation="WYAFKQWTKTYIQAPRIALGSLYKLFHQAVARILNQLDIT
NALANGLSDLLIAESPDTKVKFLESTYEYKKTNNLSTDDKKIWLAEQANKGLVI
TPPKRGKNTSYIIQASADSQAQAYKLLQYLLKNNQAVTLUSLDFGQNVNTLLVNL
NKEIIDDFQSEKLDQIAHQRLDTTAEQAGIIDYRSKSGFNOASSYKFLLEGK
LLSABLKATKAPITTPRYEVVRKQIDIEGLMRDNIQAQAYRYQMKPSEPVIRKDP
NKALILGALPGAMFAIVGTLVYATLKDRTKLD"
1645. .1650
1666. .1671
/note="putative promoter of O-antigen biosynthetic operon"
1745. .3069
/gene="wbgt"
1745. .1749
/gene="wbgt"
1756. .3069

-35_signal

-10_signal

gene

RBS

CDS

/gene="wbgt"
/note="wbgt; ORF4"
/codon_start=1
/transl_table=11
/product="putative UDP-glucose 6-dehydrogenase"
/protein_id="AAK85164.1"
/db_xref="GI:15149173"
/translation="MFYVHFMMDKMKFDTLNAKIGIIGLVGVLPLAVEFGKVVTTI
GFDINKSIDELRNHSTLBCSNLELLEATKLYACSLDALKECNVFIIVTPTIDK
HKQDPLTLIKASSETLGRKIKGDDVYIYESTVYPGATEDCIPVYKVSGLKFNIDFF
AGYSERINPGDKHRVTNLIKVTSGTPDVAEYVDQYKLIITVTKHASKIKVAEA
AKVIENTORDVNIALINELSIIFNKLIGDITLLEVAAGTKWFLPFGCLVGHGICGV
DPXYLTHKAQSGVGHPEMLAGRLNDSMGQYVWSQYKMLKQRIQVEGANVLMGL
TFKCNCPDLNRTKVIDIISLKEYNINIDIDPWCSTDEAQTGILTCEDCFKVNHYD
AIIIAVAHNEFREMGESALRALGKDEHVLFDLYVLDKRSIDMRL"
3150. .4187
/gene="wbgu"
3150. .4187
/gene="wbgu"
/note="wbgu; ORF5"
/codon_start=1
/transl_table=11
/product="putative UDP-glucose 4-epimerase"
/protein_id="AAK85165.1"
/db_xref="GI:15149174"
/translation="MDIYMSRYEETQQLIFSPKWTLITGVAGTIGSNLLEKLLKLNQ
VWIGDNFSTGHQYNDLDEYKTLVSTEQSRFCFIEGDINDLTTCQVMKGVDRVLHQA
ALGSPRSIVDITTNATNITGFLNLHAAKNAQVSFTYAASSYTGDPALPKVEE
NIGNPLSPVATKYVNEIYAQVYATYGFKTIGLRYFNVRGRDPNGAYAAVIPKWT
RAMLKGDVYINGDGETSRFCYIDNQMNTLSALAKSDKNLYNNVAGDRTLINE
LSGYIDELNLHHIDKLSIKYREFRSGDVRHSQADVTKAIDLLKYPNIKIREGLRL
SNPVYRFLKG"
4266. .5556
/gene="wzx"
4266. .4269
/gene="wzx"
4276. .5556
/gene="wzx"
/note="ORF6"
/codon_start=1
/transl_table=11
/product="putative repeat unit transporter"
/protein_id="AAK85167.1"
/db_xref="GI:15149176"
/translation="MFRKMDAGGTFLLKAIFQIGVYVFTHVSDITTFGIISYVFTV
WVFLNFDYGFRTLVKIDSDNSYSASSELLSRSGVKTYVFFFIIFPMFYSYVSDS
ISLTLVYISSAYFVCISGRFSLLQAVGRFCELYINITYIIVIGCNLFSLFTEP
LYVSALIFVYSILLVFSHKNVCPCHIKRPSILVYKDFLDATPFATLLVLLNVLS
SIDLFLKEYFSYNSVAIQVYRVNTGLIIVNVIYVLLSFYSYLLKNSWGNIRK
LQRYISLLVLLCLCYFFGIYFVGLFGDEYKVISSATFLIMFMALIKYNFWLINEL
YLVCSGNOSERVKSYCIGVISMAVFFYFIPRYGWSGAVFGSAIATLVIGIYIISVK
KDCGKILHDKYSLMMIFVPFIFFIINGQORLLY"
5614. .6797
/gene="wzy"
5614. .5617
/gene="wzy"
5625. .6797
/gene="wzy"
/note="ORF7"
/codon_start=1
/transl_table=11
/product="putative polysaccharide polymerase"
/protein_id="AAK85166.1"
/db_xref="GI:15149175"
/translation="MLIYLYPVLLLFNLIPVFPYGMNSDLERFCVPIGPIPLIFY
FEVLTSTITLRFHVSLLTKLLEGLIYLIQMLLSADISGCVLLSFFSNFI
ALVLVSFCIGKDELYLHSHVSNVNMWCMFGIICGVVKLFICYSDSNFIVLNRNAT
AIIYVCFYGVSYFYGRKSNVSVSYLFFELFSDRAGIISFISFLYQLQITKK
EKLLISLFFVPLLTIGISFTDICTRLERMLSSQVIFSGNTLTSQNDYRRVELVFI
GVDVLKENYLTIGLVANVKAIDKFLGSTNFGLAHNFYLSAQGLIGIFILLIS
VFYIMLSPIFKCGGYIGKCVFALAFYVFEYVILTPAIYIYISIFLSVVFIRNSK"
6885. .7948
/insertion_seq="IS630"

repeat_region

```

RBS      6885. 6887
CDS      6894. 7925
        /note="ORF8: Tpsae; smilar to rs630 53g ORF5 and HW38
        /transposase"
        /codon_start=1
        /transl_table=11
        /product="transposase"
        /protein_id="AAK85168.1"
        /db_xref="GI:15149177"
        /translation="MPTIAPISDERRLMQKAIHKTHDKNYARRLTAMLHGRDVRV
        DVARTLCARSSVGRWNWFTQSGVEGLSKLPAGRARWPFEEHICTLLRELKHSXGD
        FGVRSSSTELLAIKINEITGCOLNASTVRWLPAGIWRAPPTLIRDPKHDEK
        MAIHKALDCSAHEPVFYEVDVHNLKIGAWOLRQQRKVPFGQNEKAYILAGA
        LHSGTGRVCGVGNSSSALFISLLARLKATYRRAKTIITLIVDNYIIHKSRTOQWLK
        ENPKFRVYQVYSPVMNHVRLWQALHDTITRNHQCSSMWOLLKKVRHFMETVSPFP
        GGRHGLAKV"
        gene      7947. 9202
        /gene="wbgv"
RBS      7947. 7950
        /gene="wbgv"
CDS      7938. 9202
        /gene="wbgv"
        /note="ORF9"
        /codon_start=1
        /transl_table=11
        /product="wbgv"
        /protein_id="AAK85169.1"
        /db_xref="GI:15149178"
        /translation="MLLEYVERKISLAKYKVRDVIKFFVLYIASLPGIILNKNT
        VIOQKIVEISIDSESFYCYDHPMSNGRVLFPSSAFKRPKPKVKYISICVK
        DLLNKKVLYDTPAFNWOOGSLRMWIDDDNIIFENDYENNGIISVYSLSMKVYIKKI
        NPYIDVNNYKATLDFSLAKYDSYGYNNKSFSDISINLTGGTGLEFLSDEM
        LKTNKCNIDVHVNHFMAFDPGRSVMFIHRYTPKGRERLHWNLINDNVRVLI
        NESIISHCWNDEIIFGFAGIEISLNTYRLSIESCNTEKLFDFARKYSQGHPTIVH
        NRYIISDTPDKNRKIKFLFYDLVKNDYRELGLFYESMFFSYSCDLHPLRISVDNRE
        LFYDSVSGRKKRKYFMRSGICE"
        gene      9181. 10181
        /gene="wbgw"
RBS      9181. 9184

Query Match      23.4%; Score 348.8; DB 1; Length 17986;
Best Local Similarity 93.1%; Pred. No. 2.5e-92;
Matches 376; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY      4 CAGTCGGAGATGAA-AGCACACTGTGTGTACCCCATCAGCGTGTCTCCGAGGCCAGT 62
Db      16534 CAGTCGGGTGATGAACAGCAGCACCTGTGTATACCCCATCAGTGTGTCTCCGAGGCCAGT 16593
QY      63 ATTTTGTGCACACTCAATGACTACCGGAGCAGTGAACCTTCCGGTGTCTTCTCCAGC 122
Db      16594 ATTTTGTGCACGACTCAATGACTACCGGAGCAGTGAACCTTCCGGTGTCTTCCAGC 16653
QY      123 CAGTTAAGCCAGCGGTTTCCCTGCTGAAATAATGTCGGCAAAACGGGGAAGCATCAGAAGG 182
Db      16654 CAGTTAAGCCAGCGGTTTCCCTGCTGAAAGATATCGCAAAACGGGGAAGCATCAGAAGG 16713
QY      183 CGCGGGGAACCTCGTCGGGCGAGTGAACCGTGCACACTCCGGGCGAGTACATCGCGCGG 242
Db      16714 CGCGGGGCACTCGTCGGGCGAGTGAACCGTGCACACTCCGGGCGAGTACATCGCGCGG 16773
QY      243 CGCTGATACCGGCAAGATGTTGCGCAAACTCCCGCTCCGTCCGTCCAGCGGCGTATTTCAGGAT 302
Db      16774 CGCTGATACCGGAAGATGTTGCGCAAACTCCCGCTCCGTCCGTCCAGCGGCGGATTTCCGGAT 16833
QY      303 ACCCTTCGTATCAACACGCTACAAACGAGAAGACAGCTTTTGTGTTCTGATCATCCACAA 362
Db      16834 ACCCTTCGTATCAACACGCTACAAACGAGAAGACAGCTTTTGTGTTCTGATCATCCACAA 16893
QY      363 AGAAGGATATTCAGGTCGCGCAGCACTCAACGCACTCAACGCACTCAACGCACTCAACGCACT 406
Db      16894 AGAAGGATATTCAGGTCGCGCAGCACTCAACGCACTCAACGCACTCAACGCACTCAACGCACT 16937
```

```

RESULT 16
LOCUS   AP001510/c
DEFINITION
ACCESSION AP001510 BA000004
VERSION   AP001510.1 GI:10173440
KEYWORDS
SOURCE   Bacillus halodurans DNA.
ORGANISM Bacillus halodurans
REFERENCE
AUTHORS   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
1 (sites)
TAKAMI,H., Nakasone,K., HIRAMA,C., TAKAKI,Y., MASUI,N., FUJI,F.,
NAKAMURA,Y. and INOUE,A.
TITLE     An improved physical and genetic map of the genome of alkaliphilic
Bacillus sp. C-125
JOURNAL   Extremophiles 3 (1), 21-28 (1999)
MEDLINE   99184645
PUBMED    10086841
REFERENCE
AUTHORS   2 (sites)
TAKAMI,H. and HORIKOSHI,K.
TITLE     Reidentification of facultatively alkaliphilic Bacillus sp. C-125
to Bacillus halodurans
JOURNAL   Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
AUTHORS   3 (sites)
TAKAMI,H., Nakasone,K., Ogasawara,N., HIRAMA,C., Nakamura,Y.,
MASUI,N., FUJI,F., TAKAKI,Y., INOUE,A. and HORIKOSHI,K.
TITLE     Sequencing of three lambda clones from the genome of alkaliphilic
Bacillus sp. strain C-125
JOURNAL   Extremophiles 3 (1), 29-34 (1999)
MEDLINE   99184646
PUBMED    10086842
REFERENCE
AUTHORS   4 (sites)
TAKAMI,H., Takaki,Y., Nakasone,K., HIRAMA,C., Inoue,A. and
HORIKOSHI,K.
TITLE     Sequence analysis of a 32-kb region including the major ribosomal
protein gene clusters from alkaliphilic Bacillus sp. strain C-125
Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)
JOURNAL   Extremophiles 3 (3), 227-233 (1999)
MEDLINE   99411980
PUBMED    10484179
REFERENCE
AUTHORS   6 (sites)
TAKAMI,H.
TITLE     Genome analysis of facultatively alkaliphilic Bacillus halodurans
C-125
JOURNAL   (in) Extremophiles in deep-sea environments (Ed.);
: HORIKOSHI, K. TSUJII;
: 249-284; Springer-verlag (1999)
REFERENCE
AUTHORS   7 (sites)
TAKAMI,H., Masui,N., Nakasone,K. and HORIKOSHI,K.
TITLE     Replication origin region of the chromosome of alkaliphilic
Bacillus halodurans C-125
JOURNAL   Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)
MEDLINE   99356711
PUBMED    10427704
REFERENCE
AUTHORS   8 (sites)
TAKAMI,H. and HORIKOSHI,K.
TITLE     Analysis of the genome of an alkaliphilic Bacillus strain from an
industrial point of view
JOURNAL   Extremophiles 4 (2), 99-108 (2000)
MEDLINE   20263314
PUBMED    10805564
REFERENCE
AUTHORS   9 (sites)
NAKASONE,K., MASUI,N., TAKAKI,Y., SASAKI,R., MAENO,G., Sakiyama,T.,
HIRAMA,C., FUJI,F. and TAKAMI,H.
TITLE     Characterization and comparative study of the rrm operons of
alkaliphilic Bacillus halodurans C-125
```

**JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS**

Extremophiles 4 (4), 209-214 (2000)

20426005

10/972189

10 (sites)

Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N.,

Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N., Kuhara, S. and

Horikoshi, K.

Complete genome sequence of the alkaliphilic bacterium Bacillus

halodurans and genomic sequence comparison with Bacillus subtilis

Nucleic Acids Res. 28 (21), 4317-4331 (2000)

20512582

11058132

11 (bases 1 to 302150)

Takami, H. and Takaki, Y.

Direct Submission

Submitted (22-MAR-2000) Hideto Takami, Japan Marine Science and

Technology Center, Deep-sea Microorganisms Research Group; 2-15

Natsushima, Yokosuka, Kanagawa 237-0061, Japan

(E-mail: takami@jamstec.go.jp,

URL: http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html,

Tel: 81-468-67-3895, Fax: 81-468-66-6364)

Location/Qualifiers

1. 302150

/organism="Bacillus halodurans"

/db_xref="taxon:86665"

/note="alkaliphile"

23. .778

/gene="BH0826"

/gene="BH0826"

/codon_start=1

/transl_table=11

/product="transcriptional repressor"

/protein_id="BAB04545.1"

/db_xref="GI:10173441"

/translations="MLTIERHQIVLQVEKVTIHDLYEYTNASESTIRRDLELE

KQKLRHGASVLRREEPTMAETKVNNEKKKIAELASFEVDGDCFLDAGT

TVYELPIKIKIVVTNGLSDISALLEAGVETVIGGVKGTAFVGRHAIQSILQ

TFREDKAFINGVHLEDGFTDPPQAYVKETALSLSREAFMLVDHRSRGETTFATV

APIHAATVTSQVEQVIOAYETQAVKVVTL"

775. .1695

/gene="frub"

775. .1695

/gene="frub"

/note="BH0827"

/codon_start=1

/transl_table=11

/product="fructose 1-phosphate kinase"

/protein_id="BAB04546.1"

/db_xref="GI:10173442"

/translation="MIYVTINPSIDYIYOVENPQQGVNRSERDRKQPGKGINVS

VLKRLGHETKALGFLGFTGAYYRNALKEEIGLSFIEVGDTRINVKIKGQETELN

GTAPLKKEHVALLQLETELEKGVLAGSVQPMQPTIYRSMTQIAKERGAFAV

DTSGEALHEVLAAPKPSFIPKNNHSELVSKPIASIEDAIPHVORLIGEGIESILVSF

ADGALFASAEHGFHVNVPSEVRNSVGAGDSVVGAFALQESGLEDAVPAVAG

SATAFSDGCTREVERLQQLARTIKKEG"

1701. .3578

/gene="frua"

1701. .3578

/gene="frua"

/note="BH0828"

/codon_start=1

/transl_table=11

/product="PTS system, fructose-specific enzyme II, BC

component"

/protein_id="BAB04547.1"

/db_xref="GI:10173443"

/translation="MLKISELLKDDTWMLNLRASKEAVIDELVRLDLKAGRLNDAQ

FKRAIISERQSQTGVSGEIAIPHAKTAKVQPIAFGRSDAGIDYSLDGQPSHLFF

MLAASGEANHELETLSRFTFLMDETFRSLTLMQSEDEILAAIDKKEATAGEAEE

KQGVYELLAVTQPTGIATYMAADNLKSKAQLGYSIKVETNGSGVKVRLTDEIS

AKAIIIVAADTKVEMDRFHGKPVIOVPVTDGIRRPKELDQALAGKAPVYEGQAQSG

EDSAGGGRKLFYKHLNMGVSNMLPFVYVGGGILIAISPFKIKAFDPSDPSYHPFA

gene

CDS

EMLMTTGGGNAFGLMIPVLAFTAMSIADRPFGAAGMIGGLIAISTGAGPFGGLIAGF
LGSYVALGVKKVLANLPQTLGDGKITLIFYVFNIFTGIMTLMVLVIGFLAAFNFTGLQDM
LGSYGTANNVILGVILGGMAVDMGGPINKAFTFGIAMIDAGNFPGHAAFNAGGWPV
PLGIALATTLFKKKFTKQEREAGTNYILGASFTTEGAIPFAAADPCRVTPSIIVGSA
FAGGLTALFNWTLSPHGGAFVIFIGNVNNPLLYLVAITAGSIVTALLGLFWKKDAK
L"
complement(3677. .4078)
/gene="BH0829"
complement(3677. .4078)
/gene="BH0829"
/note="BH0829
unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB04548.1"
/db_xref="GI:10173444"
/translation="MARVNVMYRTVPKISGKSLNDTPSGSIRKVIHGLDDMSVSH
VKNAVQAVYQGPCKSTGHVNPVTEELKAKAERSTKMKWRGDLVNSNKTLPRLRLH
LRKENVHIELRLVALTAQHKEYSDGSSLE"
4694. .5254
/gene="BH0830"
4694. .5254
/gene="BH0830"
/note="BH0830
unknown conserved protein in B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="BAB04549.1"
/db_xref="GI:10173445"
/translation="MASSRLILIEAIMSGLAFVLSFKVIGVFWAGGGSLSLWMLPI
VMAFRGWKAGIVTGLVGLLKGIVVHPVOLVDYDPLPFTLIGLASLFAVKGSP
EGRFVAAIMIGGVLFASFRLFLCHFTSGIVMFGEYAEPTVALYSFYNISVYLPEA
LSIIVLALFGKKYSQMFQPKELGV"
complement(5283. .7682)
/gene="vpr"
complement(5283. .7682)
/gene="vpr"
/EC_number="3.4.21.-"
/note="BH0831"
/codon_start=1
/transl_table=11
/product="minor extracellular serine protease"
/protein_id="BAB04550.1"
/db_xref="GI:10173446"

gene

CDS

EMLMTTGGGNAFGLMIPVLAFTAMSIADRPFGAAGMIGGLIAISTGAGPFGGLIAGF
LGSYVALGVKKVLANLPQTLGDGKITLIFYVFNIFTGIMTLMVLVIGFLAAFNFTGLQDM
LGSYGTANNVILGVILGGMAVDMGGPINKAFTFGIAMIDAGNFPGHAAFNAGGWPV
PLGIALATTLFKKKFTKQEREAGTNYILGASFTTEGAIPFAAADPCRVTPSIIVGSA
FAGGLTALFNWTLSPHGGAFVIFIGNVNNPLLYLVAITAGSIVTALLGLFWKKDAK
L"
complement(3677. .4078)
/gene="BH0829"
complement(3677. .4078)
/gene="BH0829"
/note="BH0829
unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB04548.1"
/db_xref="GI:10173444"
/translation="MARVNVMYRTVPKISGKSLNDTPSGSIRKVIHGLDDMSVSH
VKNAVQAVYQGPCKSTGHVNPVTEELKAKAERSTKMKWRGDLVNSNKTLPRLRLH
LRKENVHIELRLVALTAQHKEYSDGSSLE"
4694. .5254
/gene="BH0830"
4694. .5254
/gene="BH0830"
/note="BH0830
unknown conserved protein in B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="BAB04549.1"
/db_xref="GI:10173445"
/translation="MASSRLILIEAIMSGLAFVLSFKVIGVFWAGGGSLSLWMLPI
VMAFRGWKAGIVTGLVGLLKGIVVHPVOLVDYDPLPFTLIGLASLFAVKGSP
EGRFVAAIMIGGVLFASFRLFLCHFTSGIVMFGEYAEPTVALYSFYNISVYLPEA
LSIIVLALFGKKYSQMFQPKELGV"
complement(5283. .7682)
/gene="vpr"
complement(5283. .7682)
/gene="vpr"
/EC_number="3.4.21.-"
/note="BH0831"
/codon_start=1
/transl_table=11
/product="minor extracellular serine protease"
/protein_id="BAB04550.1"
/db_xref="GI:10173446"

gene

CDS

EMLMTTGGGNAFGLMIPVLAFTAMSIADRPFGAAGMIGGLIAISTGAGPFGGLIAGF
LGSYVALGVKKVLANLPQTLGDGKITLIFYVFNIFTGIMTLMVLVIGFLAAFNFTGLQDM
LGSYGTANNVILGVILGGMAVDMGGPINKAFTFGIAMIDAGNFPGHAAFNAGGWPV
PLGIALATTLFKKKFTKQEREAGTNYILGASFTTEGAIPFAAADPCRVTPSIIVGSA
FAGGLTALFNWTLSPHGGAFVIFIGNVNNPLLYLVAITAGSIVTALLGLFWKKDAK
L"
complement(3677. .4078)
/gene="BH0829"
complement(3677. .4078)
/gene="BH0829"
/note="BH0829
unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB04548.1"
/db_xref="GI:10173444"
/translation="MARVNVMYRTVPKISGKSLNDTPSGSIRKVIHGLDDMSVSH
VKNAVQAVYQGPCKSTGHVNPVTEELKAKAERSTKMKWRGDLVNSNKTLPRLRLH
LRKENVHIELRLVALTAQHKEYSDGSSLE"
4694. .5254
/gene="BH0830"
4694. .5254
/gene="BH0830"
/note="BH0830
unknown conserved protein in B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="BAB04549.1"
/db_xref="GI:10173445"
/translation="MASSRLILIEAIMSGLAFVLSFKVIGVFWAGGGSLSLWMLPI
VMAFRGWKAGIVTGLVGLLKGIVVHPVOLVDYDPLPFTLIGLASLFAVKGSP
EGRFVAAIMIGGVLFASFRLFLCHFTSGIVMFGEYAEPTVALYSFYNISVYLPEA
LSIIVLALFGKKYSQMFQPKELGV"
complement(5283. .7682)
/gene="vpr"
complement(5283. .7682)
/gene="vpr"
/EC_number="3.4.21.-"
/note="BH0831"
/codon_start=1
/transl_table=11
/product="minor extracellular serine protease"
/protein_id="BAB04550.1"
/db_xref="GI:10173446"

gene

CDS

EMLMTTGGGNAFGLMIPVLAFTAMSIADRPFGAAGMIGGLIAISTGAGPFGGLIAGF
LGSYVALGVKKVLANLPQTLGDGKITLIFYVFNIFTGIMTLMVLVIGFLAAFNFTGLQDM
LGSYGTANNVILGVILGGMAVDMGGPINKAFTFGIAMIDAGNFPGHAAFNAGGWPV
PLGIALATTLFKKKFTKQEREAGTNYILGASFTTEGAIPFAAADPCRVTPSIIVGSA
FAGGLTALFNWTLSPHGGAFVIFIGNVNNPLLYLVAITAGSIVTALLGLFWKKDAK
L"
complement(3677. .4078)
/gene="BH0829"
complement(3677. .4078)
/gene="BH0829"
/note="BH0829
unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB04548.1"
/db_xref="GI:10173444"
/translation="MARVNVMYRTVPKISGKSLNDTPSGSIRKVIHGLDDMSVSH
VKNAVQAVYQGPCKSTGHVNPVTEELKAKAERSTKMKWRGDLVNSNKTLPRLRLH
LRKENVHIELRLVALTAQHKEYSDGSSLE"
4694. .5254
/gene="BH0830"
4694. .5254
/gene="BH0830"
/note="BH0830
unknown conserved protein in B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="BAB04549.1"
/db_xref="GI:10173445"
/translation="MASSRLILIEAIMSGLAFVLSFKVIGVFWAGGGSLSLWMLPI
VMAFRGWKAGIVTGLVGLLKGIVVHPVOLVDYDPLPFTLIGLASLFAVKGSP
EGRFVAAIMIGGVLFASFRLFLCHFTSGIVMFGEYAEPTVALYSFYNISVYLPEA
LSIIVLALFGKKYSQMFQPKELGV"
complement(5283. .7682)
/gene="vpr"
complement(5283. .7682)
/gene="vpr"
/EC_number="3.4.21.-"
/note="BH0831"
/codon_start=1
/transl_table=11
/product="minor extracellular serine protease"
/protein_id="BAB04550.1"
/db_xref="GI:10173446"

Query Match

Best Local Similarity

Matches 559; Conservative

0; Mismatches 304; Indels 12; Gaps 3;

Score 334.6; DB 1; Length 302150;

Pred. No. 8.9e-88;

621 ACTCAAAATTTCTACTATCAGAAACACTGGATTTAACTCTCTGAGATTACACAGCCT 680

83843 ACAATAACGACTGTGGCCAAACGCGTTAAACTTAATAATTTCTCTGAGCATGACAAA 83784

681 GAATCAATCCCTGGGGGCGTGAATTTGATATATCCACACAGATTTCAACAGCTGGATATG 740

83783 AAATCAAAACCACTGGGAGAGAATTTGATTTGATTCGAAGAATTTCTCAAGCTAGATTAT 83724

741 GAGCTCTGAAAAAGATATCAAAAGTTTCTGACAACTTCCCAAGGATTCGCCCTCGCG 800

83723 GATGCTCTTAACAAGACGTCGGCGACCTAATGAGACAGCTCAAGATTGTCGCTCGCT 83664

QY	801	GATTATGTCATTATGTCCTCTTTATTCGTCATGCTTGGCACGCTGCGGAAACATAC	860
Db	83663	GACITTTGGTCATTATGTCCTCTTTATTCGTCATGCTTGGCACGCTGCGGAAACATAC	83604
QY	861	AGGACATATGATGCGCGGAGCGCGCAGTGGTGGTCAGCAACGTTTGAACCGCTGAAC	920
Db	83603	CGTATAGTCATGCGCGGAGCGCGCAGTGGTGGTCAGCAACGCTTGGCCCTCTTAAT	83544
QY	921	AGCTGGCGGATACGTTATCTGATTAAGACCGCTGATGCTGTGGCCAGCAAGAA	980
Db	83543	AGCTGGCGGATACGTTATCTGATTAAGACCGCTGATGCTGTGGCCAGCAAGAA	83484
QY	981	AAATACGCTCCATCTTCTGCGGAGACCTGATGCTCTGACTGTAATGTTGCCCTT	1040
Db	83483	AAATACGCTCCATCTTCTGCGGAGACCTGATGCTCTGACTGTAATGTTGCCCTT	83424
QY	1041	GAATCCATGGAATTAAGCGCTGGGATTTGCTGGCGGAGAGAGAACTGGGAGTCG	1100
Db	83423	GAATCCATGGAATTAAGCGCTGGGATTTGCTGGCGGAGAGAGAACTGGGAGTCG	83364
QY	1101	GACCTGG---TATACCTGGGCGCTGACAAAGCGCTCTTCAGATACCGGGATAAAAC	1157
Db	83363	GAAGAGGATATTACTGGGCGGAGCGAAAGAGTGGTTAACGGGTGATAACCTTACTCA	83304
QY	1158	GGGA-----AATTCAGAAACCTCTTGGCGGCCACGACAGATGGACTTATTTATGCAAT	1211
Db	83303	GGTATGCTGAGCTTGAGATCCACTTTCAGCTCTTTCAGATGGGCTTATTTACGTTAAC	83244
QY	1212	CCTGAAGCGCCCGTGGAAACACAGATCTCTGCTTCCGCGAAGATATCAGGAGCT	1271
Db	83243	CCTGAAGCTCCAGACGGTAAACACAGATCTCTGCTTCCGCGAAGATATCAGGAGCT	83184
QY	1272	TTTTACGATGGCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACA	1331
Db	83183	TTTTGCGCGGATGGAATGAACGATGAAGAACGCTTCCCTTATCGCTGCGGACATACA	83124
QY	1332	TTTGTTAAAGACATGTTGTCAGCTCTCTCGAAATATGATTTGGCGAGCGGCTGATGCT	1391
Db	83123	TTTGTTAAAGCTCAGCTGTCAGGAAATCCAGATCA---TGTGTCAGAGCGGAGAGCT	83067
QY	1392	GCACCTGTGGAGGAGCGGAGCTGGGATGGAAATAAATGTTGTCAGGAAACGGCAAA	1451
Db	83066	GCTCGGATGAGGCACAGGCTTAGGCTGGCAAAACACATACGCTTGGCAAGGTCGC	83007
QY	1452	TATACCATCAGCTGCGCTGGAAGCGCTGGCTGTC 1486	
Db	83006	GACAGCATCATGTCGCTTGAAGCGGCTTGGAC 82972	
RESULT 17			
AF486647			
LOCUS			
DEFINITION	AF486647 3289 bp DNA linear BCT 25-MAR-2002		
ACCESSION	Rhizobium etli plasmid pCFN42f katG regulator (oxyR) and katG		
VERSION	(katG) genes, complete cds.		
KEYWORDS	AF486647.1 GI:19702287		
SOURCE			
ORGANISM	Rhizobium etli.		
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;		
AUTHORS	Rhizobium etli		
TITLE	Rhizobiaceae; Rhizobium.		
JOURNAL	1 (bases 1 to 3289)		
REFERENCE	Vargas,C., Encarnacion,S., Davalos,A., Reyes,A., Mora,Y.,		
AUTHORS	Garcia,A., Brom,S. and Mora,J.		
TITLE	Only one catalase, katG, has been detected in Rhizobium etli, this		
JOURNAL	gene and the regulator OxyR are encoded by a plasmid replicon		
REFERENCE	unpublished		
AUTHORS	2 (bases 1 to 3289)		
TITLE	Vargas,C., Encarnacion,S., Davalos,A., Reyes,A., Mora,Y.,		
JOURNAL	Garcia,A., Brom,S. and Mora,J.		
Submitted (22-FEB-2002)	Metabolic Engineering, CIPN-UNAM, Avenida		
Unversidad s/n, Cuernavaca, Morelos 6210, Mexico			

Location/Qualifiers	
1..3289	
/organism="Rhizobium etli"	
/db_xref="taxon:29449"	
/plasmid="pCFN42f"	
complement(6..947)	
/gene="oxyR"	
complement(6..932)	
/gene="oxyR"	
/codon_start=1	
/transl_table=11	
/product="katG regulator"	
/protein_id="AAL93240.1"	
/db_xref="GI:19702287"	
/translation="NKNLTAKLYFEALRGDFRRRAAGSCAVSOPALSMQIKLEQ	
ELGSELFPERSREVLTLPFGQTFALRVDRILRAVDELGEFARASDSFLTRGLIIP	
LTAPILLPLVHARLNKSPAGIQIEVRETHIKGLVHELGAQQLDMALVAPLVSESLPTE	
LSLFNRRIRAGPAPKRPVPERALREMLRLLEEGHCFRQDALSCKVGAARPRE	
IMEGSLSTLVQMVAGIGITLIPEMAVPVETRAHVSISRPPSPAPRTTIGMWLNF	
HPMGNHLPWKLAKARSRLVWRKTSPSAINGG"	
complement(942..947)	
/gene="oxyR"	
1064..3260	
/gene="katG"	
1064..1070	
/gene="katG"	
1077..3260	
/gene="katG"	
/codon_start=1	
/transl_table=11	
/product="katG"	
/protein_id="AAL93241.1"	
/db_xref="GI:19702289"	
/translation="MDNPTDTAGKCPVAHKNRPGNSNRDMWNPQNLVQIILHNSGRA	
DPLGKDFDAEDIOEARSRAEKLHALMTSDWMPADFGHYGLFIRMAHWSAGTY	
RUTDGGGACGHQHPAPLNSHPDNALDKARLLWPKQKYNRISHWDLILLGNV	
ALDSMGFTGTFAGGCADWPEPELIRARMANDEETVALIAGGHTFGTHGAGDPSFIGAE	
VPEGPNGPDVAAARDIETLARMANDEETVALIAGGHTFGTHGAGDPSFIGAE	
PEGGAIEDGGLKSKSFGVGKDAITAGLEVTSQTPTKWSNYEFENLFAVNEWELTK	
SPAGAHQWAKNAEASIPDAYEPFKKHVPTMLTDLSLRFDPIYEKISRFLNPDQF	
ADAFARAKFLTRDMGPKRVYFGPELPAEDLIWDVTPAVDPHPVDKDIKAEKAKV	
LATQTLVQELSTAWASASTFRGSDKRGANGARILAPOKDEANOPALAKLVGL	
EGIQDFNAOTGAKKISLADLILVLAGAGYKAAAGNNAVSVPLTFRMDASBQAT	
DAHSFAPLRIDGFFNYNVRKQFMPKEMLVDRAQLLITLTPMTIVLVGLRVLK	
AGNPEHVSQRAEPTVNDFFVNLDDVGQWYPAPEKGLYRPRKTRGAAKWTGTRVD	
LIFGSHSQLRAFEVYQADAKQKFKVDFVAANVMNADREDLR"	
BASE COUNT 657 a 1051 c 1011 g 570 t	
ORIGIN	
Query Match 22.4%; Score 333.6; DB 1; Length 3289;	
Best Local Similarity 66.8%; Pred. No. 5.8e-88;	
Matches 537; Conservative 0; Mismatches 259; Indels 8; Gaps 4;	
QY 688 ATCCCTGGGCGCTGATTTTGTATGATGCC-ACCAGATTTCAACAGCTGGATGGAGGCT 746	
Db 1210 ATCCGCTCGTAAGGACTTCGATATGCCGAGATATCCAAAGCTCGATCTCGACGG 1269	
QY 747 CTGAAAAAGATATCAAGATTTGTCGACAACTTCCAGGATTTGTCGCCCTCGGATTA 806	
Db 1270 CTGAAAAAGACTTC-ACGCGCTGATGACGAGTTCGACGACTGTGGCGCGGCTTC 1328	
QY 807 GGTCAATATGGTCTTTCTTTTATTCGTATGGCTTGGCAGCGTGGCGGACATCAGGACA 866	
Db 1329 GGTCAATATGGCGGCTCTTCATCCGATGGCTGGCAGCGCGGCGGCTATCGCATC 1388	
QY 867 TATGATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 926	
Db 1389 ACCGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1448	
QY 927 CCGGATAACGTTAATCTGGATAAAGCCCTCGATTCGATTCGTCGCCAGTCAAGAAAAATAC 986	
Db 1449 CCGGACAAATGCCAACCTTGACAAGGCGCGCGCTCTCTTTGGCCGATCAAGCAGAAATAC 1508	

QY 987 GGCTCCAGTATTCTCGGCGAGACCTGATGGTCTGACTGGTAATGTTGCCCTTGAATCC 1046
Db 1509 GGCACCGCATCTCTCGGCGCGACCTTCTGCTCCTACCGGCAACGTCGCGCTCGAGTCC 1568
QY 1047 ATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAAATGACTGGGAGTGGACCTG 1106
Db 1569 ATGGGCTTTAAGACGCTCGGCTTCGCGCGGCGCTGCGCCACCTCTGGGAGCGGAAGAG 1628
QY 1107 GTATACCTGGGGCTGACACACAGCCTCTTGGCAGATACCGGGATAAACG---GGA 1163
Db 1629 CTCTACTGGGGGCTCTGAAGGCACCTGGCTCGGCGAGGAGCGCTATAGCGCGAAGCAT 1688
QY 1164 CTTACAGAACTCTTGGCGCCACGACGATGGGACTTATTTATCTCAATCTTGAAGCCCC 1223
Db 1689 CTGGCGAACCCGCTTGGCGCGCTGCGAGATGGGCTCATCTACCTCAATCCCGAAGCCCG 1748
QY 1224 GTGGAAACACAGATCCTCTGGCTTCCGCGAAAGATATACGGAAGCTTTTTCACGCTATG 1283
Db 1749 AATGGCAATCTCAGCCGCTGCTGCGAGCGCGGACATTCGCGAACCCTTGGCCCGCATG 1808
QY 1284 GCCATGGATGATGAGGACTGTGGCCCTGATCGCGGAGGCGATACATTTGTTAAAGCA 1343
Db 1809 GCGATGAACGACGAGGAACCGTGGCACTGATCGCGGCGGTCTACCTTCGCGCAAGACG 1868
QY 1344 CATGGTGCACGCTCTCTGAAATATGATTTGGCGCAGGCGCTGATGGTGCACCTTGGAG 1403
Db 1869 CATGGCGCGCGCATCC---GTCTTCATCGCGCGGCAACCGGAAAGCGCGCGATCGAG 1925
QY 1404 GACGAGGACTGGGATGGAATAAATAATGTGTTACAGGAACCGCAAAATATACCATCAC 1463
Db 1926 GACCAGGCTCGGCTGGAAGAGCTCTTTTCGCGACCGCGGTGCGCAAGGACGCGCATACC 1985
QY 1464 AGTGGCTGGAAGGACCTGGTCTG 1487
Db 1986 GCGGCGCTCGAGGTTACCTGGTCTG 2009

RESULT 18
AE007314/c
LOCUS
DEFINITION
Sinorhizobium meliloti plasmid pSymA section 120 of 121 of the complete plasmid sequence.
ACCESSION
AE007314 AE006469
VERSION
AE007314.1 GI:14524456
KEYWORDS
Sinorhizobium meliloti.
SOURCE
Sinorhizobium meliloti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Sinorhizobium.
REFERENCE
1 (bases 1 to 10578)
Barnett,M.J., Fisher,R.F., Jones,T., Komp,C., Abola,A.P., Barloy-Hubier,F., Bowser,L., Capela,D., Galibert,F., Gouzy,J., Gurjal,M., Hong,A., Huizar,L., Hyman,R.W., Kahn,D., Kahn,M.L., Kalman,S., Keating,D.H., Palm,C., Peck,M.C., Surzycki,R., Wells,D.H., Yeh,K.-C., Davis,R.W., Federspiel,N.A. and Long,S.R. Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti pSymA megaplasmid
Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9883-9888 (2001)
21396509
11481432
2 (bases 1 to 10578)
Barnett,M.J., Fisher,R.F., Jones,T., Komp,C., Abola,A.P., Barloy-Hubier,F., Bowser,L., Capela,D., Galibert,F., Gouzy,J., Gurjal,M., Hong,A., Huizar,L., Hyman,R.W., Kahn,D., Kahn,M.L., Kalman,S., Keating,D.H., Palm,C., Peck,M.C., Surzycki,R., Wells,D.H., Yeh,K.-C., Davis,R.W., Federspiel,N.A. and Long,S.R. Direct Submission
Submitted (29-MAR-2001) Biological Sciences, Stanford University, 371 Serra Mall, Stanford, CA 94305, USA
Location/Qualifiers
1. .10578
/organism="Sinorhizobium meliloti"
/strain="1021"

repeat_region

/db_xref="taxon:382"
/plasmid="pSymA"
420. .1478
/standard_name="ISrm2011-2/ISrm11"
/note="Sma3018; predicted by homology"
/evidence-not_experimental
/rpt_family="ISrm2011-2/ISrm11"
516. .923
/gene="Sma2373"
516. .923
/gene="Sma2373"
/codon_start=1
/transl_table=11
/product="Tm2011-2a transposase"
/protein_id="AAK65941.1"
/db_xref="GI:14524457"
/translation="MARPSNDLRERVDAVTGEGLSRAAAKRFGIGISTADWYRR
FRETGSAAQGMCGHPRKLSGPHRAWLRCRERDFTLHGLVAELSERGLKVDYRAV
WTFVHEGSLSYKARRWSPANGSGPTSPATGHGQ"
989. .1462
/gene="Sma2375"
-989. .1462
/gene="Sma2375"
/function="Elements of external origin; Transposon-related functions"
/codon_start=1
/transl_table=11
/product="Tm2011-2b transposase"
/protein_id="AAK65942.1"
/db_xref="GI:14524458"
/translation="MAPLRGWAPRGLRVGYAFGHWTMTFVAALRADRVSAFFILD
GPINGFRFYVQVVLPELKAGDIVILDNLSGHKQGEIRAAIRKAGARLFFPKYSP
DLNPIEKLFKIKHLEAQAQRSDAIHDELRIHQAVTPOECAVFKKAGYERA"
1664. .2833
/gene="Sma2377"
1664. .2833
/gene="Sma2377"
/function="Cell processes; Transport of small molecules"
/note="glimmer prediction; similar to putative chloramphenicol transporter in several organisms, including Salmonella typhimurium, Afl18107, Klebsiella pneumoniae, AAF27726, and Enterobacter aerogenes, AAD22144"
/codon_start=1
/transl_table=11
/product="putative transmembrane transport protein"
/protein_id="AAK65943.1"
/db_xref="GI:14524459"
/translation="MTVNVSPVTVMFLLSIFGTVLGIAGTDLVLPAIPAMPTALGGTAA
LAQMLAAGGTLVGLTFGLGARYSRKLLVSLGLFAVTSLISAYAPTLEMLVI
LRFAGAGGPAVFPAGFIHGLYPGDKAPSMFGRSGTESITPALAPAGAYLMTVG
GQTSFLMLAGLAILCAVGSWAYRQSLPDLLEAVHOSYMSIIRNGDFLRHLSAL
SLGSLILFFGAPAVMTGALMTIGDFILQVFGIALFILASNASLARRFIERMI
MIGTSLVLGFLILLDITSLGSRSLVPLMWTANGAFGIRGPIGFHQAIVASRDH
SNGAALVAAIIGITAGGTAAAPFIVNGWPLAASSLAALLILCLKLIGSTA"
complement(2924. .5107)
/gene="Sma2379"
complement(2924. .5107)
/gene="Sma2379"
/EC_number="1.11.1.6"
/function="Cell Processes; Protection responses; detoxification"
/note="glimmer prediction; highly similar to catalase/peroxidase from Streptomyces reticuli, CAA74698"
/codon_start=1
/transl_table=11
/product="catalase/peroxidase"
/protein_id="AAK65944.1"
/db_xref="GI:14524460"
/translation="MDQKSDSAGKCPVAHTAPGRSNRDNWPDQLDVQVLRHRSGLSD
PLGNFTNAEEFKLLDALKRDLRALMTDSQDWWPADFGHYGGGLFIRNAWHISAGTYR
ITDGGGAGGQGOORVFNWPDNANLDRLLWPIKQYGNRISWADLLITLGNVA
LESMGFKTFGAGRVDMWPELEFWGPEGTWLGDERYSGERQLSEPLAAVQMGLIYV

VGGEORSVLVVPTPTGSGWIDPAAMESLEVQLONGDVASVAVOY SYLASWLSLLVEPDY
GADTAQALFRVDYDWRLPKNSRPRLYLHGLSLGAMNSERSTDMDYVDYDFQGLM
SGPPYKSLWRSFTRDSCFSAWLPFRNGSLVRFMNDQGVAAADPFVWGPRIYVL
QXASDVFDFDNDLYREPDWIKPPRPDVSPELRYPPVTFLLQTLDMAMATTAPIG
HGVTPAGDYPVGVGIEVANVOGWSLQOIAELKRLHRLP*
7907. .8404
/gene="mir6724"
7907. .8404
/gene="mir6724"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52962.1"
/db_xref="GI:14026365"
/translation="MSDLVVIYVPEAKAEVRORLPALQKEYLINISDAVIATKTEA
GNKLMQIMTIVAGVSGFWGLLIGVIFLNPVIGVATGAASGALGALTDVGINDA
FKKLSJATLQPNAAFLVLIHGMTADKVLISAIRKGGVVLKTSLDTEKREQLREALSQ
ASEAA"

gene

CDS

9371. .10375
/gene="mir6726"
9371. .10375
/gene="mir6726"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52963.1"
/db_xref="GI:14026366"
/translation="WVNIATGILVLAALYFARSLIMPLAFSLFVIALAWPIQMALQR
RIPRLAALITLLVVIIVIGSAIVGSLGELISQWLFTYNTGRLOAIYASWIDWLEGH
GLAIVSPGLSELSIRGLVLSLTRLNSVAGFAVLTFYVMLGLEAEDLQRRRS
SLAIVPOGEMLHGRIAGKIRREVMVRSIACILGLVVAFTLVAGLELASAWAIA
FALNVPFGPLMATIPLFLAIAQSGSNQALIVFVGLVIOFALGSLIEPLFGAS
LAISPAVFAVFFSFMVGVPFAF IGVPILIVVVTWCAHEQSRWIRIATLLSGRST
*
10712. .11839
/gene="mir6727"
10712. .11839
/gene="mir6727"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB52964.1"
/db_xref="GI:14026367"

gene

CDS

/translation="MANAPAGNNAHAILEPPANPAEPTLVVPEOTPIPCFEFTSGD
LPPDPAWRNSFATMLDPFPGDTEAGFAGTQVINDGLDALAGVSVNDLFTSLA
GVRKDPDHLVTLIQDGRSOTIIPSKSGFDAGSVVHPLGRVPEGSVTDSEMLLL
FYPRDCRMSGLAAAEFSLRGGMGPLADYLASLVQRLPTMGVNDLFGILSATRA
MLQVCIAPSPDHLTEAHDPISATLLERARRFVQHILFQPDLSIEOMTREIGTSRSLY
RLFEASGGIVVHYIOHRLLAHAALADPNDRRLILDIADEYGFGDGAETSFRAFREFG
YCPSEVRTGVKNGPSHWHMDEVETAAPGLRLQQLRLRQAQ*
complement(11950. .12498)
/gene="ml16729"
complement(11950. .12498)
/gene="ml16729"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52965.1"
/db_xref="GI:14026368"

gene

CDS

/translation="MTAYENGTVSSLPPTSVRIILGLVILVGGIFVLGLALATVISA
IVIGVTAIVVGAEMHAFHFMKGWGFIMQFLGLILYVYVFGAMVROPVYSGALVTV
IGLLVVASGVRLRVGVWRWSVGVMTWLSGTFGLAGLVLTGWPSGLWVGLGLLG
IDLISHGMALWLTGWLPTAKTA*
complement(12495. .15662)
/gene="ml16730"
complement(12495. .15662)
/gene="ml16730"
/codon_start=1
/transl_table=11
/product="multidrug-efflux transport protein"

gene

CDS

Query Match 21.7%; Score 323.2; DB 1; Length 340857;
Best Local Similarity 64.6%; Pred. No. 2.4e-84;

Matches 532; Conservative 0; Mismatches 283; Indels 9; Gaps 3;
QY 672 CACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTTGAATATCCACCAGATTTCACAG 731
DB 197459 CAGTCTCTCTGTGCCAGCCGATGGGTGAGCGCTTCGACTATCCCGAGGAATCAAGAGC 197518
QY 732 CTGGATATGGAGGCTCTGAAAGAAAGATATCAAGAGATTTGCTGACAACTTCCAGGATTGG 791
DB 197519 CTCGACCTCGACGCCCTCATCAAGGACCTGACACGCTGATGACGATTCGAGGATGG 197578
QY 792 TGCCCTGCGGATTATGTCATATTGTCCTTTCTTTATTCTGATGCTTGGTCCACGGTGCC 851
DB 197579 TGGCCGCGCGATTTCGCGCCACTAGCCCGCTGTTTCATCGGATGCGCTGGCAGACGCCA 197638
QY 852 GGAACATACAGGACATATGATGCCGGGGAGGCCAGTGGTGTGAGTGTGAGTGTGAGTGTGAA 911
DB 197639 GGCACCTACCGCATTCGCCAGCGCCGCGCGCGCGCGCTGGCCAGCAGCGCTTTCGCG 197698
QY 912 CGCTGAACAGCTGCGCGGATACGTTAATCTGGATTAAGCCCGTCTGATTCGTGTGCCCA 971
DB 197699 CGCTCAACAGCTGCGCGGACAAATGCAACCTCGACAAGCCCGCGCGCTTTTGTGCGCG 197758
QY 972 GTCAAGAAAAATACGCTCCACTATTTCCTGGGGAGACCTGATGTCCTGACTGGTGAAT 1031
DB 197759 ATCAACAGAAATATGCGCCCAAGATCTCTGCGCGGACCTCTGATCTCACCGGAC 197818
QY 1032 GTTGCCTTGAATCCATGCGGATTTAAACGCTGGGATTTGCTGCGCGAAGAGATGAC 1091
DB 197819 GTCGCGCTGGAATCGATGGCTTCAAGACCTTTGGCTTTCGCGCGCGCGCGAGGTC 197878
QY 1092 TGGGAGCTCGGACCTG---GTATACTGGGCGCTTGACAAACAGCCTCTTTCAGATATACCGG 1148
DB 197879 TGGGAGCGCTGAGCAGGACGCTTTACTGGGCTCCGAAAGGCAAGTGGCTGGCGCGCAGCGC 197938
QY 1149 GATA---AAAAAGGAAACTTTCAGAAACCTTTCGCGCCACGACGATGGGACTATTAT 1205
DB 197939 TACAGCGCGGACCGTGCACCTGCAGAACCCGCTGCGCGCTGCAGATGGGCTGATCTAC 197998
QY 1206 GTCATCTCTGAAGCCCGCGGTGAAACACAGATCTCTTGGCTTCCCGGAAAGATATCAG 1265
DB 197999 GTCAATCCGAGGAGCGAAGCGCAATCCGGATCCGCTGGCGCGCGCGCGATATCAG 198058
QY 1266 GAAGCTTTTTCAGTATGCCATGATGATGAGAGACTGTGCCCTGATCGCGGAGGG 1325
DB 198059 GACACATTCGCGGTATGGCGATGAACGAGGAGAAACCGTAGCACTCATCTCGCGCGCG 198118
QY 1326 CATACATTTGGTAAAGACATGTCGAGCGCTCTCTGAAATATGTTTGGCGCAGGCGCT 1385
DB 198119 CATACGTTCCGCAAGACCCATGCT---GCGGGTGACGCGAGCCTGGTGGGTGTCGAGCG 198175
QY 1386 GATGGTGACCTGTGGAGGAGCAGGACTGGGATGGAATAAATAAATGTGTGTACAGGAAC 1445
DB 198176 GAAGCGCTGATATCGAGCAGCAGGCGCTTGGCTGGCGGAGCAATTCGSCACCGCAAG 198235
QY 1446 GCGAATATACCATCACCAGTGCCCTGGAAGAGCGCTTGGTCGAC 1489
DB 198236 GCGGTGACGCCCATCGGACGCGCTGGAAGATCATTTTGGACGAC 198279

RESULT 21

AL646061/c

LOCUS

DEFINITION

ACCESSTION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AL646061 198050 bp DNA linear BCT 07-DEC-2001
Ralstonia solanacearum GMI1000 chromosome, complete sequence;
segment 5/19.
AL646061 AL646052
AL646061.1 GI:17427781
Ralstonia solanacearum.
Ralstonia solanacearum
Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
Ralstonia.
1 (bases 1 to 198050)
Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,

Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
Chandler,M., Choisme,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schlex,T.,
Siguer,P., Thebaud,P., Whalen,M., Wincker,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.

Genome sequence of the plant pathogen *Ralstonia solanacearum*

Unpublished
2 (bases 1 to 198050)

Boucher,C.A.

Direct Submission

JOURNAL

TITLE

AUTHORS

JOURNAL

Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Bodu, 75010 Paris, France, ILMG CNRS
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
F31326 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
<http://sequence.toulouse.inra.fr/R.solanacearum.html>.

COMMENT

FEATURES

source

1. .198050
Location/Qualifiers

/organism="Ralstonia solanacearum"

/strain="GM1000"

/db_xref="taxon:305"

151. .642

/gene="RSC0770"

/note="RS05084"

151. .642

/gene="RSC0770"

/EC_number="1.-.-.-"

/function="small molecule metabolism"

/note="Product confidence : putative

Gene name confidence : hypothetical

predicted by Codon usage

predicted by Frameb"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="PUTATIVE TRANSMEMBRANE DEHYDROGENASE (SMALL

SUBUNIT) OXIDOREDUCTASE PROTEIN"

/protein_id="CAD14472.1"

/db_xref="GI:17427782"

/translation="MHTRPSSGDDPAGLTRQWLQSLAATAASLVGSVLYIYAAQ

APGEPLDPMGLSQTALRSALDRVAGRLAALQASAGFAAQSLAAGTLS

GLQGETALRIQAWYTGVDYGVVYEQALMYGVVSDILVIRTYCPNRRPGFWAEPPEV

KQS"

642. .2261

/gene="RSC0771"

/note="RS05083"

642. .2261

/gene="RSC0771"

/EC_number="1.-.-.-"

/function="small molecule metabolism"

/note="Product confidence : putative

Gene name confidence : hypothetical

predicted by Codon usage

predicted by Homology

predicted by Frameb"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="PUTATIVE TRANSMEMBRANE DEHYDROGENASE (LARGE

SUBUNIT) OXIDOREDUCTASE PROTEIN"

/protein_id="CAD14473.1"

/db_xref="GI:17427783"

/translation="MAETQAEQADIVVVGSGVAGALYAYELARACKSVLMLEAGPRL

PRWEIREFRQADKMDFNAPYPTFPWPHPEYGDNNYLIILKGHKFDSQITRAVGG

TWHAASTWRLPNDLKLRSYGIARDWPIQYDLERYGRAEALGWGPNDEBLG

SPRSQYPMAPLPSFNERTVKDALNAHPALRVYTEPVARNRPYDGRPTCCGNNNC

MPICPTGAMINGIVHVEKAEQAGARLIENAVVYKLEVGAKRIVAAHYKDPKGVDRHY

gene

CDS

EGKWFVLAANGIETPKMLMSTGADFFPKGVGNRSDVMGRNLMHPGTCVSVYADRKLM
PGRQEMTSLIGRDEGFRATQAGKIHLSNIRIEQETTRIFKAGKLLKPAELDAR
IRDOAARVQVDSFEHILPENRIVPSATCTDALGIPRPEITRIDDYVKRSVAVHR
EYVTAARVLGTVQVDFHDFAPNNHITGATMGSDPKDSVVDKDCRTDFHPLNFSS
SATWPTGCTVNVTLTIAALRLADIQLKKEA"
2263. .3552
/gene="RSC0772"
/note="RS05082"
2263. .3552
/gene="RSC0772"
/EC_number="1.-.-.-"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Frameb"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE OXIDOREDUCTASE DEHYDROGENASE
(CYTOCHROME C SUBUNIT) SIGNAL PEPTIDE PROTEIN"
/protein_id="CAD14474.1"
/db_xref="GI:17427784"
/translation="MRAKAAASFLLSVLGAAAAARDADAPADAAQIKRGEYLAIGD
CVACHTAPGKPFAGGLPQIPLMLGTIYSSNTPDKTGKWRYEERAVRQGVSD
DGENLYPAMPYPSYAKINADMDRLYAFVRYGVAOVAPPPTIRWPLNMRWPLKLM
NLAFLKRAPYLPKGDPTNWRGALYLAQGLAHACCTCHPRGFAMOEKAMDERCKGYLA
GSTLAGNSAYNITSDPYSGIGGWPEQIYQVLRGSGVPGIQAAGGAEAVQHSFSRM
TERDYRALEYLRTPAVSTGNSARHGWTGTPATDIALRGKPIETTDIARLYLGN
ATCHQADRGTPDGYIPLLHNSAVGARDTNLVQVMLHGHTERRAGRHIGMPAFGRQ
LSDAALANYVTRQFGDPATPALTADEIAKRMPQ"
complement(3559. .5781)
/gene="RSC0773"
/note="RS05081"
complement(3559. .5781)
/gene="RSC0773"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Frameb"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD14475.1"
/db_xref="GI:17427785"
/translation="MRTSARVILGITPIVIVAAAVFVITFDWNRKAPYLNDHVSO
AVGRPFAINGDLALTWKRPESGWRAYVPMPLIASDIAIGNPDWARQPHVATVQRLL
TFVLEALPLLAHRIVIPSVTLDTPMIALERDSQNRNNTWTFHGAQDSQSDWKLELE
IAFSGSVQLDDEVKIRHLAATLDVDNQALYTAASGTAIRGSEPAAPSAGAAASTA
PAAPDPOQYGIATVKGTYNNAIRGSRAGGVLRLOQDARRYPLOADVTGKTRID
LAGTLTPSSLTALDLRLHLSGSAWHLPLTVGLPDTPTDFTRGLIGELRQGS
WRYEQFTGRVGGSDLGTLTFAMRPECPQLTGLVLSHQLLFADLAPIIGADSNAS
KQRDAPVRQADKVLPEVPEPTRNAIDADVKFTGERIVRTADLPIDHLVTHIKQ
DAVTLDPNFGVAGGTLTNLRDQSGAPMQASAAALAHFKIRQLFPNIESMHASI
GEINGDARLSATGNSVAALIGASNGELKILVEQGVSKFILEAMGLNVGNVILTKLFG
DKOVSINCAAGDEAVSDLAQARTFVVDODVIDVTGATSFKHESLDFTHPDSKGL
RVFSRLPLYKGYTKHPDYSVNPVVALRAGGAVALAPAPVNAVLPVLEQLQPDUS
PCGKLADVVRORPTAPPPGKVRSGQTRAGDGSPPDPAAPARQAPALPLRDPTTAGG
"
5883. .6785
/gene="RSC0774"
/note="RS05080"
5883. .6785
/gene="RSC0774"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon usage

Db	8622	GACCTCATGATCTCAGGGCAACGTCGCGCGCTGGAGTCGATGGGCTTCAAGACCTTCGCG	8566
Qy	1068	TTTTCGTCGGCGAAGAAAGATGACTGGGAGTCGGAAGCTG---GTATACTGGGGGCTGAC	1124
Db	8562	TTTCGCGCGCGCGCAAGGATGATGGGAACCGAAGAGGATGCTCTACTGGGTTCGGA	8503
Qy	1125	AACAAGCCTCTTCAGATACACCGGGATAAAAAG---GGAACTTCAGAAACCTCTGTC	1181
Db	8502	ACCACCTGGCTGGCGGACCAAGCGCTACACCGCGGATCGCGATCTGGAGAAACCGCTCGCC	8443
Qy	1182	GCCACGCAGATGGGACTTATTTATGTCAATCTCAAGGCCCGCTGGGAACACAGATCCT	1241
Db	8442	CGCGTCAGATGGGCTGATCTACGTCAATCCGGAAGGCCGAACCGCAACCCCGATCCA	8383
Qy	1242	CTGCGTCCGCGAAGATATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAG	1301
Db	8382	ATCGCGGCGCCAGGACATCCCGAGACGTTCCGCCGATGCGATGAACGATGAAGAG	8323
Qy	1302	ACTGTGGCCCTGATCGCGGAGGGCATACATTTGGTTAAAGCACATGGTCAGCGCTCTCT	1361
Db	8322	ACGTCGCGCTGATCGCGGCGGCCACTCTCTCGGAAGACCCACGCGGCGCTCGGCT	8263
Qy	1362	GAATAATGATTGCGCAGGCGCTGATGGTGACCTGTGGAGAGCAGGACTCGGATGG	1421
Db	8262	CCAACGTG----GGCGCGGAGCGGAAGCGCGCATCGAAGAACAGGCGCTGGCTGG	8207
Qy	1422	AAAAATAATGTGTTACAGAGAAACCGCAATATACCATCACAGTGGCCCTGGAAGAGCC	1481
Db	8206	AGCAGCGCTTCGCGACCGCGCAGGCGACCGAGCGCCATCACAGCGCTTGGAGTGACC	8147
Qy	1482	TGTCGAC 1489	
Db	8146	TGGACCAC 8139	
RESULT 22			
SRFRPS			
LOCUS			
DEFINITION	Streptomyces reticuli furS DNA	linear	BCT 05-MAR-1999
ACCESSION	Y14317		
VERSION	Y14317.1	GI:3758888	
KEYWORDS	catalase/peroxidase; cpeB gene; Fe regulatory protein; furS gene.		
SOURCE	Streptomyces reticuli.		
ORGANISM	Streptomyces reticuli.		
REFERENCE	Bacteria; Firmicutes; Actinobacteriae; Actinobacteridae;		
AUTHORS	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.		
	1. (bases 1 to 2904)		
	Zou, P.-J., Borovok, I., de Orue Lucana, D.O., Mueller, D. and Schremf, H.		
TITLE	The mycelium-associated Streptomyces reticuli catalase-peroxidase, its gene and regulation by FurS		
JOURNAL	Microbiol. 145, 549-559 (1999)		
REFERENCE	2. (bases 1 to 2904)		
AUTHORS	Borovok, I.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUL-1997) I. Borovok, Angewandte Genetik der Mikroorganismen, FB Biologie/Chemie, University of Osnabrueck, Barbarastrasse 11, Osnabrueck D-49069, FRG		
REMARK	revised by submitter 12-Oct-1998		
COMMENT	Related entry: Y14336.		
FEATURES	Location/Qualifiers		
source	1..2904		
	/organism="Streptomyces reticuli"		
	/strain="Tu45, described by H.Zahner, University of Tuebingen"		
	/db_xref="taxon:1926"		
	/clone="pUBB1"		
	/clone_lib="BamHI minilibrary"		
gene	78..533		
	/gene="furS"		
RBS	78..82		
	/gene="furS"		
	/note="putative"		

SOURCE	Streptomyces reticuli.
ORGANISM	Streptomyces reticuli Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE	1 (bases 1 to 2904)
AUTHORS	Zou, P.-J., Borovok, I., de Orue Lucana, D.O., Mueller, D. and Schremf, H.
TITLE	The mycelium-associated Streptomyces reticuli catalase-peroxidase, its gene and regulation by Furs
JOURNAL	Microbiol. 145, 549-559 (1999)
AUTHORS	2 (bases 1 to 2904)
TITLE	Borovok, I.
JOURNAL	Direct Submission Submitted (23-JUL-1997) I. Borovok, Angewandte Genetik der Mikroorganismen, FB Biologie/Chemie, University of Osnabrueck, Barbarastrasse 11, Osnabrueck D-49069, FRG
REMARK	revised by submitter 12-Oct-1998
COMMENT	Related entry: Y14336.
FEATURES	Location/Qualifiers 1. .2904 /organism="Streptomyces reticuli" /strain="Tu45, described by H.Zahner, University of Tuebingen" /db_xref="taxon:1926" /clone="pUBB1" /clone_lib="BamHI minilibrary"
gene	78. .533 /gene="furs"
RBS	78. .82 /gene="furs" /note="putative"

QY 924 TGGCCGGAATACGTTAATCTGATGAAGCCGCTCGATTCCTGTGGCCAGTCAAGAAAAA 983
|||||
Db 5782 TGGCCGGAATACGTTAATCTGATGAAGCCGCTCGCTTGTGGCCATCAACAGAA 5841
|||||
QY 984 TACGGCTCAGTATTTCTCTGGGAGACGTCGATGGTCTGACTGGTAACTTGCCTTGAA 1043
|||||
Db 5842 TATGTCAGAAATCTCTGGGCGGACCTGTTATCTCTCGGGTAACGTGGCGTAGAA 5901
|||||
QY 1044 TCCATGGGATTTAAACGCTGGGATTTGTCGCGGAAGAAGATGACTGGGAGTCGGAC 1103
|||||
Db 5902 AACTCGGCTTCGCTACCTTCGTTTGGTGGCGGTCGTAAGACGCTCTGGGAACCGAT 5961
|||||
QY 1104 CTGGTATATCTGGGCGCTCACACAGCCTCTTGACAGTAACCGGATAAACGGGAA 1163
|||||
Db 5962 CTGGATGTTAATCTGGGCTGATGAAAAAGCCTGGCTGACTCACCGTCATCCGAAGCGCTG 6021
|||||
QY 1164 CTTCAAAAACCTCTTGGCCCGCCAGCAGATGGGACTTATTTATGTCATCTCTGAAGCCGCC 1223
|||||
Db 6022 GCGAAGCACCGCTGGTGCACCGAGATGGTCTGATTTAGTTAACCCGGAAGCCCG 6081
|||||
QY 1224 GTGGGAACACAGATCTCTGCTTCGCGGAAGATATACAGGGAAGCTTTTTCACGTATG 1283
|||||
Db 6082 GATCAGCGCGGCAACCGCTTCTCGCGCAGCAGCTATCCGCGACCTTCGGCAACATG 6141
|||||
QY 1284 GCCATGGATGATGAGGAGACTGTGCCCTTGATCGCGGAGGAGCATACATTTGGTAAAGCA 1343
|||||
Db 6142 GGCATGAACGGAAGAAACCGTGGCGCTGATTTGGCGGTGGTATACGCTGGGTAAACCC 6201
|||||
QY 1344 CATGTGTCAGCTCTCTCAAAAATGATTTGGCGCAGGCGCTGATGGTCACCTGTGGAG 1403
|||||
Db 6202 CACGGTGGC---GTCCGACATCAATGTAGTCTGATCAAGAGCTGCACCGATGAA 6258
|||||
QY 1404 GAGCAGGACTGGGATGGAATAAATGTTGACAGGAACCGCAATATACCATCAC 1463
|||||
Db 6259 GAACAAGGTTAGTTGGCGAGCACTACCGCAGCGCGGTTGGCGCAGATGCCATTACC 6318
|||||
QY 1464 ACTGGCTGGAAGGACCTGTC 1486
|||||
Db 6319 TCTGGTCTGGAAGTAGTCTGGAC 6341
|||||

RESULT 25

ECOW87 96484 bp DNA linear BCT 12-AUG-2002
E. coli chromosomal region from 87.2 to 89.2 minutes.
L19201
L19201.1 GI:304961
23S ribosomal RNA; 5,10 methylenetetrahydrofolate reductase; 50S
ribosomal protein L31; 5S ribosomal RNA; 6-phosphofructokinase;
CDP-diglyceride hydrolase; DNA polymerase I; L-rhamnose isomerase;
Spot 42 RNA; aspartokinase II/homoserine; atp21II; catalase HPI;
catalase hydroperoxidase I; cdh gene; chemosensory transducer; cpxA
gene; cystathionine; cyrB gene; dehydrogenase II; dsbA gene;
extragenic palindromic element; fdhB gene; fdhE gene; fdg gene;
fdhH gene; fdoI gene; formate dehydrogenase-O alpha subunit;
formate dehydrogenase-O beta subunit; formate dehydrogenase-O
gamma; gamma-synthase; glnA gene; glng gene; glnL gene; glpF gene;
glpK gene; glpX gene; glutamine synthetase; glycerol kinase; hslU
gene; hslV gene; katG gene; kdsT gene; manganese superoxide
dismutase; metB gene; metF aporepressor; metF gene; metJ gene; metL
gene; mob gene; msrA gene; periplasmic sulphate binding protein;
pfkA gene; polA gene; priA gene; primosomal protein replication
factor; rhaA gene; rhaB gene; rhaD gene; rhaE gene; rhaS gene; rhaT
gene; rhamnose permease; rhamnulokinase; rhamnulose kinase;
rhamnulose-1-phosphate aldolase; rpmE gene; rfa gene; rrlA gene;
sbp gene; sodA gene; spf gene; suppressor; tpiA gene;
triosephosphate isomerase; triosephosphate mutase.
Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 96484)

SOURCE

ORGANISM

Plunkett, G. III, Burland, V., Daniels, D.L. and Blattner, F.R.
Analysis of the Escherichia coli genome. III. DNA sequence of the
region from 87.2 to 89.2 minutes
Nucleic Acids Res. 21 (15), 3391-3398 (1993)
93347969
8346018
This sequence was determined as part of the E. coli Genome Project
(Frederick R. Blattner, director) at the University of
Wisconsin-Madison. Supported by award HG00301 from the NIH Human
Genome Project. A preliminary report was presented at the Small
Genomes meeting, March 28-30 1993, Paris, France. The entire
sequence was independently determined from E. coli MG1655; overlaps
and conflicts with other sequence determinations are annotated. The
start of this entry overlaps the end of the entry ECOW850 (M87049)
by the three bases of an EcoRI half-site. Data kindly submitted in
computer readable form by:
Guy Plunkett III
Laboratory of Genetics
University of Wisconsin
445 Henry Mall
Madison WI 53706
USA
Phone: 608-262-2534
ecoli@genetics.wisc.edu
Email: 608-263-7459.
Fax: Location/Qualifiers
1..96484
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"
/cissue_lib="lambda"
/note="This sequence comprises the following lambda
clones: EC21-16, EC27-1225, EC27-649, EC19-119, EC23A58,
EC27-267, EC18-126. M13mp19 or Janus vectors were used for
subcloning."
1..19281
/note="corresponds to lambda clone EC21-16"
1..2063
/gene="rrlA"
<1..2063
/gene="rrlA"
/product="23S ribosomal RNA"
/note="CG Site No. 203; 23S rRNA"
1..3
/gene="rrlA"
/note="Eco RI half-site is 3bp overlap with end of M87049
ECOW850"
2157..2276
/gene="trfA"
2157..2276
/gene="trfA"
/product="5S ribosomal RNA"
/note="CG Site No. 210; 5S rRNA"
2282..2321
/note="putative"
2459..2483
/note="putative"
2504..2540
/standard_name="REP; repetitive extragenic palindromic
element."
/note="contains 1 REP sequence"
complement(2546..3058)
/standard_name="ORF_f170"
/note="matches PS00017: ATP_GTP_A"
/codon_start=1
/transl_table=11
/protein_id="AAB02991.1"
/db_xref="GI:304962"
/translation="MIPLAFAAWSGTGKTLKKLPALCARGIRPGLIKTHHMD
VDKPKDSYELKAGAAOTIVASQORWALMTETPDEEELDQFLASMDYTSKLDLILY
EGFKHEETAKIVLFRDGAHRPEELVIDRHVAVASDVPLNDVALLINDVEGLADF
VVEWMQKQNG"

CDS

repeat_region

terminator

terminator

rRNA

gene

misc_feature

rRNA

gene

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

rRNA

misc_feature

gene

complement(3055. .3639)
 /gene="mob"
 complement(3055. .3639)
 /gene="mob"
 function="biosynthesis of molybdopterin"
 /note="CG Site No. 921"
 /codon_start=1
 /transl_table=11
 /protein_id="AAB02992.1"
 /db_xref="GI:304963"
 /translation="NNLTTTGVVLGGKARRMGVYDKLLELNGKPLWQHVDALM
 TOLSHVNNRHOEYQASGLKVLIEDSLADYPCPLAGMLSYMQOEAGEWFLFCPCDT
 PYPIDLAARLNHQKRDAPVWVWVHGERDHPHIALVNRALIEPLLLLEYLQAGERRVWV
 MRLAGHAVDFSDHKDAFVNVNTPPELAKWQKR"
 complement(3642. .3668)
 /note="55% match to consensus; resembles
 flhC/flhD-regulated promoter; putative"
 3653. .3680
 /note="promoter matrix score of 48; putative"
 complement(3691. .3720)
 /note="promoter matrix score of 66; putative"
 3709. .3978
 /standard_name="ORF_o89"
 /codon_start=1
 /transl_table=11
 /protein_id="AAB02993.1"
 /db_xref="GI:304964"
 /translation="MKCKRLNEVIELLPQAWKEPDPLNLQFLQKLAKESGPDGELAD
 LDDILYHLKMRDSAKDAVIFGLQKDYEDFTALLRANGVIKE"
 3743. .3898
 /note="Predicted bend of 84.96 degrees"
 3848. .3875
 /note="promoter matrix score of 69; putative"
 4056. .5042
 /standard_name="ORF_o328"
 /codon_start=1
 /transl_table=11
 /protein_id="AAB02994.1"
 /db_xref="GI:304965"
 /translation="MNSAFTFOTLHPDITMDALPEHGRVDSGLPLNSYENRVYOF
 QEDRRFRVVKYRPERWADQILIEHQFALQLVNDEVPVAPVAFNGFTLLNHQGY
 FAVPSVGGQFADNDQMEAGVTRIGRHQYGRKQLRIHRTIGLNEYLIEPRKLF
 EDATLPSGLKAAFLKATDELTAATVAHWRDEPTVLRHGDCHAGNLRWDRPMFVDL
 DARNPAGVQDLWMLNGKAEQRMQLETHIEAYEFSEFDTAELIGLIEPLRAMRLVY
 YLAWLMRWADPAFPRNFWLTGEDYWLQRATTFIEQAKVLQEPPLQLTPMY"
 4126. .6322
 /note="corresponds to X63186; ECOPPPA(1. .2192)"
 misc_feature
 4173
 /note="absent in X63186; 'c' in this sequence"
 /replace=""
 misc_difference
 4481
 /note="absent in X63186; 'c' in this sequence"
 /replace=""
 misc_difference
 4792/4793
 /note="in X63186; absent in this sequence"
 /replace="a"
 misc_difference
 4865/4866
 /note="in X63186; absent in this sequence"
 /replace="c"
 misc_difference
 4883
 /note="in X63186; 't' in this sequence"
 /replace="a"
 misc_feature
 5044. .5718
 /note="corresponds to M77746; ECODSF(1. .675)"
 5059. .5685
 /gene="dsbA"
 5059. .5685
 /gene="dsbA"
 /function="disulphide bond formation in vivo"
 /codon_start=1
 /transl_table=11
 /protein_id="AAB02995.1"
 /db_xref="GI:304966"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 FEATURES
 source

Philipps-Universitaet-Marburg, Marburg, Germany
 2 (bases 1 to 3018)
 Scherer,M.
 Direct Submission
 Submitted (25-JAN-2001) Scherer M., Biochemistry,
 Max-Planck-Institut fuer Terrestrische Mikrobiologie,
 Karl-von-Frisch-Str. 35043 Marburg, GERMANY
 revised by author 10-MAY-2001

location/Qualifiers
 1. 3018
 /organism="Aspergillus nidulans"
 /db_xref="taxon:5072"
 393. 2612
 /gene="cpeA"
 393. 2612
 /gene="cpeA"
 /EC_number="1.11.1.6"
 /codon_start=1
 /product="catalase-peroxidase"
 /protein_id="CAC59821.1"
 /db_xref="GI:15384983"
 /translation="MGSNECPYSRQNTANGGGGNNRDPDLKLNLRHNSVNP
 LDRGFDYAFNSLDYFGLKRDLEALMTDSQDWMPADFGHYGLFIRMAHWSAGTVRV
 FGRGGGGGQOQRFAPLNSWPDVNSLOKARLLWPVKQYKSKISADLILAGNAL
 ESMGFTFEPAGSRDTEADQSVFQWGEKWLNDVRYLNGELNDPLAASHMLIYV
 NPGPNKNDPVLAKDIRITFGFMANDEETVALLAGHTFGKTHGAGPATHLGKEP
 HGAGIELQGWESGESGTHAITSGLVITWKTPTKWSNOFFETLYFYDWELTKS
 PAGAHQYVANGVEFFIPVDPFSPKHPHMLTDLSDRYDPEYERISRRFLENPDQFA
 DAFARAWFLTHRDVGRVPLVYQGPVSEVLVWQDPVPLDHPVIDNDIDTATLKAIL
 NSGISTDGLFSTAWASASTPRGSKRGANGARIRLSPKWNKVNQSPWLSLALE
 KIOQFNDAQSDTKRSLADLIVLAGAASLEKARDAGHNVSVFPPGRDATQEOFD
 VDSFNLEPIADGFNRYGRGTPVLTFEFLDKAQLNLNLSPPELTVLIGLRLVNNY
 DRNLGVTFRQQLNDFVNLDDGMQVMPADDTNEIFIGSDRTKQGWAKASRAD
 LVFGSHAEIRLRAISEVSSGDEAKFYKDFVAWEKRVSNLDRFLDKTGLAQRKLPQI"

BASE COUNT 660 a 883 c 789 g 686 t
 ORIGIN

Query Match 20.4%; Score 303.4; DB 8; Length 3018;
 Best Local Similarity 61.2%; Pred. No. 6.2e-79;
 Matches 528; Conservative 0; Mismatches 326; Indels 9; Gaps 2;

632 CTACTATCCAGAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 464 CTGTTGGCGGACGACTCAACTGACATCTCTGTCACACACACTCGTCTTAATCC 523
 692 CTGGGGGCTGATTTGATTATGACACAGATTTCAACACTGGATGAGGCTCTGAA 751
 524 CTTGGATAAGGCTTTGACTACTATGCGCGGTTCAATAGCTCGACTACTCGGACTGAA 583
 752 AAAAGATATCAAGATTTGCTGACAACTCCACAGGATGCTGCCCTCGGATATGCTCA 811
 584 GCGGGATCTGGAGGCACTATGACAGACTCCACAGGACTGGTGGCGGCGGACTTTGGTCA 643
 812 TTATGTCCTCTTTTATTCGATGGTGGCAGGCTGGCGGACATACAGGACATATGA 871
 644 CTATGGCGGACTCTTATCGCATGGCTGGCAGACTGCTGGACAGTATCGGCTTTGA 703
 872 TGGCCGGGGAGGCGGCTGGTGGTACGACAGCTTTTGAACCGCTGACAGCTGGCGGGA 931
 704 CGGTGGCGGGTGGCGGACAGGCTGACCAAGCTTCGCTCCGCTCAACAGCTGGCGGGA 763
 932 TAAGCTTAATCGATTAAGCCCGCTGATGCTGCTGTCGACAGTACAGAAAAAATACGCTC 991
 764 TAAGCTGACGTTGGCAAGGCTCGTCTCTTTATGGCCCATCAAGCAGAGTACGGAAG 823
 992 CAGTATTCTTGGGAGACCTGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051
 824 CAGATCTATGCGCTGATGCTGATCTTCGCGGGAATGCTGCGCTTGAATCAATGCGG 883
 1052 ATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTCGGACCTGGGTATA 1111
 884 GTTCAAGACCTTTGGGTTTGGCGGTGGCCGAGTATGATACCTGGGAGCAGACCACTGCGT 943

1112 CTGGGGCCTGACACAA-----GCCTCTTGCGAGATAACCGGATAAAAACGGGAAACT 1165
 944 CTTCTGGGAGCGGAGAAAGAAATGGTTGGTAAATGATGTCCGTACTTGAACGGAGAACT 1003
 1166 TCAGAAACCTCTTTCGCCCCACGAGATGGGACTTATTTATGTCAATCCTCAAGGCCCGCG 1225
 1004 CGACACCCGCTTGGGCAATCACATCGGTCCTTATTTAGTTTATTCAGAGGACCCAA 1063
 1226 TGGAAAAACAGATCCTCTTCCGCGAAAGATATCAGGAAAGCTTTTTCACGTATGGC 1285
 1064 CAAGAACCCGACCCGCTTCTCGCGGCAAGGATATCCGCATCACCTTTGTGCAATGGC 1123
 1286 CATGGATGATGAGAGACTGTGGCCCTGATCGGGAGGAGCATACATTGGTGAAGCACA 1345
 1124 CATGAATGATGAGAGACTGTGGCCCTGATTTGGTGGACACACAGTTCGGAAGACGCA 1183
 1346 TGGTGCAGCGCTCTCTGAAAAATGATTGGCGGAGGCGCTGATGTGTCACCTGTGGAGGA 1405
 1184 CGG---CGCGGGCCTGCAACCCATCTCGGCAAGAACCACATGTTGCGGGTATTGAGTT 1240
 1406 GCAGGACTCGGATGAAAAATAAATGTGTGTACAGGAAACGCAAAATATACCATCACCAG 1465
 1241 ACAAGGCTTAGCTGGGAGAGCGCTTCGAGTCTGGGACCGCGGACATGCTATCACCAG 1300
 1466 TGGCCTGGAAGGAGCCCTGGTCTGA 1488
 1301 CGGCTGGAGGTGATCTGGACCA 1323

RESULT 28
 AE005625
 LOCUS
 DEFINITION
 of 290.
 Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 244
 ACCESSION
 AE005625 AE005174
 AE005625.1 GI:12518851
 VERSION
 AE005625
 KEYWORDS
 Escherichia coli O157:H7 EDL933.
 Escherichia coli O157:H7 EDL933
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE
 1 (bases 1 to 10667)
 Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
 Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
 Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
 Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
 Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
 Welch,R.A. and Blattner,F.R.
 Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
 Nature 409 (6819), 529-533 (2001)
 21074935
 MEDLINE
 PUBMED
 11206551
 REFERENCE
 2 (bases 1 to 10667)
 Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
 Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
 Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
 Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
 Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
 Welch,R.A. and Blattner,F.R.
 Direct Submission
 TITLE
 Submitted (22-OCT-2000) Laboratory of Genetics, University of
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 JOURNAL
 FEATURES
 Location/Qualifiers
 1. 10667
 /organism="Escherichia coli O157:H7 EDL933"
 /strain="EDL933"
 /serotype="O157:H7"
 /db_xref="taxon:155864"
 /note="enterohemorrhagic"
 <1. .251
 /note="O-island #162: Region of the EDL933 chromosome not
 homologous to E. coli K-12 MG1655"
 complement(253. .861)

```

/gene="yix"
/notes="Z5492"
complement(253..861)
/gene="yix"
/function="orf; Unknown function"
/notes="Residues 1 to 202 of 202 are 97.52 pct identical to
residues 1 to 202 of 202 from Escherichia coli K-12 Strain
MG1655: B3937"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG59138.1"
/db_xref="GI:12518852"
/transl_table="MKRLILSLVSPAFAMQPQTGDIIFQISRSSQSKAIQLATH
SDYSGTMLVNRNKKPVYFAVGKVTPLKQWIAHGEKGVYVRRVEGLSVEQOQK
LAQAKRYLGPYDFSPSWSDROCYCEVVKVYQNALGMVGEQQLKEFDLSNPLV
QAKLKERYGNIPLEEVVSPQAFDAPQLTVAKWPLFSW"
complement(921..1238)
/gene="metJ"
/notes="Z5493"
complement(921..1238)
/gene="metJ"
/function="regulator; Amino acid biosynthesis: Methionine"
/notes="Residues 1 to 105 of 105 are 100.00 pct identical
to residues 1 to 105 of 105 from Escherichia coli K-12
Strain MG1655: B3938"
/codon_start=1
/transl_table=11
/product="repressor of all met genes but metF"
/protein_id="AAG59139.1"
/db_xref="GI:12518853"
/transl_table="MAEWSGEYISPYAEHKKSEQVKKIVTSIPLKVLKILTDERRR
QVNLRLATNSLCEALFHTAFTGQPLDDADLRKERSDEIPEAAKEIMREMGINPET
WEY"
1515..2675
/gene="metB"
/notes="Z5494"
1515..2675
/gene="metB"
/function="enzyme; Amino acid biosynthesis: Methionine"
/notes="Residues 1 to 386 of 386 are 99.48 pct identical to
residues 1 to 386 of 386 from Escherichia coli K-12 Strain
MG1655: B3939"
/codon_start=1
/transl_table=11
/product="cystathionine gamma-synthase"
/protein_id="AAG59140.1"
/db_xref="GI:12518854"
/transl_table="MTRKQATIAVRSGINDOEYGVVPPHLSSTYNTGFNEPRAH
DYSRGNPTRVQRALEGGAGVLTNTGMSAHLVTYFLKPGDLLVAHPCYV
GSLRFLDSLRKRGCTRVLDQGDQALRAALNPKLVLVESPSNPLLRVYDIKIC
HLAREVAVSVVNTFISPALQNPALGADLVLSCTKYLNHSDVVGAVVIADPDV
VTELAWANNITVGGAFDSYLLRLGLTLVPRMELAQNAKAIYVLOQVPLVKLY
HPSLPENGHIAAQQOQGFGLSFLDGDQOTLRLFLGLSLFTLAESLIGVESLI
SHAATMHAGNAPARAAGISETLRLISTIGIEDGDLIADLNGFRAANK"
2678..5110
/gene="metI"
/notes="Z5495"
2678..5110
/gene="metI"
/function="enzyme; Amino acid biosynthesis: Methionine"
/notes="Residues 1 to 810 of 810 are 99.87 pct identical to
residues 1 to 810 of 810 from Escherichia coli K-12 Strain
MG1655: B3940"
/codon_start=1
/transl_table=11
/product="aspartokinase II and homoserine dehydrogenase
II"
/protein_id="AAG59141.1"
/db_xref="GI:12518855"
/transl_table="MSVIAQAGAKGRQLHFKFGSSLDVRCYKLRVAGIMAYSQPDDM
MVSAAGSTTNOLINLWIKLSQDRLSAHOVQOTLRRYQCDLISGLLPAEADSLISAF
VSDLERLAALDSDINDAVYAEVVGHEVNSARLMSAVLNQOGLPAALWLDAREFLRAE

```

```

RAAQPOVDEGLSYPLQQLLVQHPGKRLVVTGFIISRNAGETVLLGRNGSDYSATQIG
ALAGYSRTVITSDVAGVTSADPRKVDACLPLLRIDEASELSARLAAPVTHARTQLAPV
SGSEIDQLQRCSYTPDOGSTRIVRSLAGTGARIVTSHDDVCLIFQVPTSDQFKLAH
KEIDQLKRAQVRPLAVGVNDROLLOFCYTSFVADSALKILDEAGLPGELSLRQGLA
LVAMGAVYTRNPLCHRFWQQLKQOPVEFTWQSDGSLVAVLTGPTGTELSIQHLQ
SVFRAKRLGLVLFGRGNIGRWLELFAREQSLTSARTGFEFVLAVGVDSRRSLISYD
GIDASRALAFNDENAEODEESLFLWMAHRAHPYDVLVDVLTASOQLADQYLDFAHGF
HYTSANKLAGASDNKRYQIHDAFETGRHLYNATVGAGLPINHTVDLIDSGDTIL
SISGIPSTLSWLFQFCGVPFTLVQAOQOGLTEPDRDLGKDVMRKLVILAR
EAGYNTEPDQVRVESLVPACRGGSIDHFEFENGDELNEQVORLEAAAREMGLVLRVA
REDANGKARVGYEAVREDHPLASLLPCDNVFALESRWYRDNPLVIRGPGAGRDVTAGA
IOSDNLRAQLL"
5459..6349
/gene="metF"
/notes="Z5496"
5459..6349
/gene="metF"
/function="enzyme; Central intermediary metabolism: Pool,
multipurpose conversions of intermed. met."
/notes="Residues 1 to 296 of 296 are 99.66 pct identical to
residues 1 to 296 of 296 from Escherichia coli K-12 Strain
MG1655: B3941"
/codon_start=1
/transl_table=11
/product="5,10-methylenetetrahydrofolate reductase"
/protein_id="AAG59142.1"
/db_xref="GI:12518856"
/transl_table="MNFHASORDALNQSALAEVQOINVSFEFFPPTSEMBOITLWNS
IDRLSLKPKFVSVTYGANSRDRTHSIIGIKORTGLEAAPHLCIDATPDELRTI
ARDYNNGIRHIVALRGDLPGSGKPEMYASDLTLKEVADFEDLSVARYPEVHPEAK
SAQADLLNKRKVDAGANRAITQFFDVESEYLRDRRCVSAGIDVEILPGILPVSNFK
QAKKFDLMTNVRIPAWMAQMFQDGLDDDDAETRLKLVGANIAMDVMKILSREGVKDFHYT
LNRAEWSYAICHTLGVPRGL"
6678..8858
/gene="katG"
/notes="Z5497"
6678..8858
/gene="katG"
/function="enzyme; Protection responses: Detoxification"
/notes="Residues 1 to 726 of 726 are 99.86 pct identical to
residues 1 to 726 of 726 from Escherichia coli K-12 Strain
MG1655: B3942"
/codon_start=1
/transl_table=11
/product="catalase; hydroperoxidase HPI(I)"
/protein_id="AAG59143.1"
/db_xref="GI:12518857"
/transl_table="MSTSDDIHNTTATGKCPFHGGHSDQAGAGTTTRDWPNNOLRVD
LLNQHNSRNPLGEDDYRKESKLDYGLKDKLALTESOPWMDWGSYAGLFIIR
MAWHGAGTYSIDGRGAGGQGFAPLNSWPDNVSIDKARLLPIKQYQKQISWA
DUFILAGNVALENSGFRTEFGAGREDVWEPDLVNWGDEKAWLTHRHPEALAKAPLG
ATEMGLIYVNEGPDHSGPLSAAAIATFNGMNDDEETVALLTAGHTLGTGAG
PTSNVAPDEAPAFIEQGLGWASTYSGVGADAITSGLEVWTTQPTOWSNYFFENLF
KFWVOTRSPAGAIQFVADPEIIPDPDPKRRKPTMLVDTLRFDPPEFEKISRR
FLNDQAFNEAFARAFKLTDRMDGPKSVYIGPEVPKEDLIWQDPLQPIYNTQDII
IDLKFAISGLSVSELSVAVNASASTFRGGKGGANGARLALMPQRDWDVNAAYR
ALPVELEKIQESKASLADILVLAGVGVKEAASAGLSIHVPFAPRVDARQDDQDI
EMFLELPADGFRNTYRLADLSTTESLIDKAAQLTAPETALVGMRYLGFANFD
GSKNGVTDRTVGVLSNDFVNLDMRYENKATDESEKELFEGDRDTEGVTYASRADL
VFGSNSVLRAVEVYASSDAHEKFKVDEVAWVKVNNLDRFDLL"
8919..9857
/gene="yijE"
/notes="Z5498"
8919..9857
/gene="yijE"
/function="orf; Unknown function"
/notes="Residues 1 to 312 of 312 are 98.71 pct identical to
residues 1 to 312 of 312 from Escherichia coli K-12 Strain
MG1655: B3943"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG59144.1"

```

```

/db_xref="GI:12518858"
/translation="MSGTKLHTLSLMSAAGSNPLAISGLVLTLLWSYSWIFMKQVT
SYGAFDTALRCFLGFLVILLRGRMRPTPKYLTIALALLQTCGAVGLQAWAL
VSGCGAVILSYTPFWYVFAAVFLGRGRQGYFAILIAAFGLFLVLPQWOLFDS
SMKSAMAILSGVSWGASATVKARLYARHPVDLLSTSMQLYAALVSVVALLVPO
REIDQPTVFWALAYSAILATALANSWLFLVNLKPLASTASTLAVPCGVLFMSWL
LGENPGAVEKGVILVILALALVSRKKEAVSRKI"
gene
complement(9884. .10501)
/gene="yijr"
/note="Z5499"
complement(9884. .10501)
/gene="yijp"
CDS

```

[illegible]

RESULT	29
AF027168	2885 bp DNA linear BCT 11-NOV-1997
LOCUS	Caulobacter crescentus catalase-peroxidase (katG) gene, partial cds.
DEFINITION	
ACCESSION	AF027168 GI:2605729
VERSION	
KEYWORDS	Caulobacter crescentus.
SOURCE	Caulobacter crescentus
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter
REFERENCE	1 (bases 1 to 2885)
AUTHORS	Steinman,H.M., Fareed,F. and Weinstein,L.
TITLE	Catalase-peroxidase of Caulobacter crescentus: function and role in stationary-phase survival
JOURNAL	J. Bacteriol. 179 (21), 6831-6836 (1997)
MEDLINE	98012985
PUBMED	9352936
REFERENCE	2 (bases 1 to 2885)
AUTHORS	Steinman,H.M., Fareed,F. and Weinstein,L.
TITLE	Direct Submission
JOURNAL	Submitted (26-SEP-1997) Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461-1602, USA
FEATURES	Location/Qualifiers
source	1..2885
	/organism="Caulobacter crescentus"
	/strain="CB15"
gene	/db_xref="taxon:155892"
	1404..>2885
	/gene="katG"
CDS	1404..>2885
	/gene="katG"
	/codon_start=1
	/transl_table=11
	/product="catalase-peroxidase"
	/protein_id="AAC45850.1"
	/db_xref="GI:2605730"
	/translation="METPARNTPTMDAKVEDNTAGKCPMGHGRGPANRDMWPPOSRLRE GLNWAHSRPNSMIGDEADYAEAFKSLLDLDAVVSDLHALMTDSQEWPADEFGHYGLGFIR LWHAAGTYRTITDRCGAGGGQORFAPLNSDPNTNLDKARLIWPIKQRYGAKLGSRA DLVLYGVNALESWMGFKTFEAGGRADQPEPELYWGPESTWLDDKRYSGERELDSPL GAVOMGLIYVNPPGNPNPDLASARDIRETFARMANDDETVALLAGHTTFCGKAAGA GDASLVGVEPEGGAIEAQCDFGWASKHGTCGPDATIGGPEVINTQTPTWSNHFFPNLR FYKENWTQSAPAGAKQAOKNAHPADIADFDPNKTHIVPRMLTSDDLALRDFDPATEKISR RFYENPDQFAARAWFLKTRHDMPGIPRGYLGVLVPKEELIWDDLPPIAVDHPHLADDIK DIAALKAKILATIGLSASDLVSTAWASASTYRQSKDRKGANGARI"
BASE COUNT	503 a 1027 c 923 g 432 t
ORIGIN	
Query Match	20.3%; Score 302.2; DB 1; Length 2885;
Best Local Similarity	62.0%; Pred. No. 1.4e-78;
Matches	513; Conservative 0; Mismatches 308; Indels 6; Gaps 2;
Qy	663 CTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCGTGATTGTGATTGCCACCAGA 722
Dd	
Db	1539 CTGACCAGACGCCCGCGCGCTCATCCGATGGCGGAGCGGTTCGACTACGCCGAGCGG 1598
Qy	723 TTTTCAACAGCTGGATATGGAGGCTCTGTAAAAAAGATATCAAAGATTTCGTGACAAC TTC 782
Dd	
Db	1599 TTCAAGAGCCTCGACCTGGACGCGGTGCTGCAGCGACCTGCACGCCCTGATGACCGCACGC 1658
Qy	783 CAGGATTTGGCCCTGGCGATTATGGTCATATGGTCCTTTCTTTATTTCGTATGGCTTGG 842
Dd	
Db	1659 CAGGATTTGGCCCGCTGACTTCGCGCCATATGGCGGCGCTGTTCATTCGCTTGGCTTGG 1718
Qy	843 CACGGTCCCAGAACATACAGGACATATGATGGCCGGGAGCGCGCACTGGTGTGTCAGCAA 902
Dd	
Db	1719 CACGCGCCGCGACCTATCGCATCACCGACGCGCGCGCGCGCGCGCGCGCGCGAGCAG 1778
Qy	903 CGTTTTGAACCGCTGAACAGCTGGCCGGATAACGCTTAATCTGGATAAAGCCCCGTGATTG 962
Dd	

Db	1779	CGTTCGCCCCCTGACAGCTGGCGGCAACACCAATCTCGACAGGCCGCCGCTG	1838
QY	963	CTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGAGACCTGATGTCCTG	1022
Db	1839	CTGTGGCCGATCAAGCAGAGTAGTGGCGCCCAAGCTGTCTGGCGCCGACCTCTATGTCGT	1898
QY	1023	ACTGGTATGTTGCCCTTGAATCATCGGATTTAAACACGCTGGATTGCTGGCGGAAGA	1082
Db	1899	GTCGGCAACGTCGCCCTTGGAGTCGATGGCTTCAAGACCTTCGGCTTCGGCGGCCGCG	1958
QY	1083	GAAGATGACCTGGAGTCGGACCTGGTATCTACTGGGGCCCTGACACAGCCTCTTTGCAGAT	1142
Db	1959	GCCGACCACTGGAGCCCGAAGAGCTGTACTGGGGTCCGAAAGCAGCTGGCTGGACGAC	2018
QY	1143	AACCGGGATAAAACG---GGAACATTCAGAAACCTCTTCCGCCCGCAGATGGGACTT	1199
Db	2019	AAGCGCTACAGCGCGGCGCTGAGCTGGACTCGCCCTGGCGCGGTCCAGATGGGGCTG	2078
QY	1200	ATTATGTCAATCTCTGAAGGCCCGGTGGAAACACAGATCCTCTGGCTTCGGCGAAAGAT	1259
Db	2079	ATCTACGTCATCTCCGAAGCCCGCAAGCAACCCGACCCGCTGGCTTCGGCGCGGAC	2138
QY	1260	ATCAGGGAAGCTTTTACGTATGGCCATGATGATGAGAGACTGTGGCCCTGATGCGG	1319
Db	2139	ATTCGCGAGACTTTCGCCCGGATGCCATGAATGACGAGGAGACCGTCGCGCTGATGCC	2198
QY	1320	GGAGGCGATACATTTGGTAAAGACATGTTGCACGCTCTCTGAAAAATGTTATGGCGCA	1379
Db	2199	GCGGCCACACTCTCGGCAAGCTCTACGG---CGCGGGCGACGCTCCTCTGGGGGCTT	2255
QY	1380	GGGCTGATGTGACCTCTGTGGAGGAGCAGGACTGGGATGGGAAAAATAATGTGGTACA	1439
Db	2256	GAGCCGGAAGCGCGCGGATCGAGGCCCGAGGGCTTGGTGGCGGAGCAAGCAGCGACG	2315
QY	1440	GGAACCGCAATATACCATCACCAGTGGCTTGAAGGAGCTGCTGTC	1486
Db	2316	GGCAAGGCGCGCGGACCGCATCACCAGCGCGCGCGGAAGTGTCTGGAC	2362
RESULT 30			
AE005967			
LOCUS	AE005967	10677 bp	DNA linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus CB15 section 293 of 359 of the complete genome.		
ACCESSION	AE005967		
VERSION	AE005967.1	GI:13424686	
KEYWORDS	Caulobacter crescentus CB15.		
SOURCE	Caulobacter crescentus CB15		
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter.		
REFERENCE	1 (bases 1 to 10677)		
AUTHORS	Nierman,W.C., Feldblyum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolony,J.F., Smit,J., Craven,M.B., Khouri,H., Shetty,J., Berry,K., Utterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and Fraser,C.M.		
TITLE	Complete genome sequence of Caulobacter crescentus		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)		
MEDLINE	21173698		
PUBMED	11259647		
REFERENCE	2 (bases 1 to 10677)		
AUTHORS	Nierman,W.C., Feldblyum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Laub,M.T., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolony,J.F., Smit,J., Craven,M., Khouri,H., Shetty,J., Berry,K., Utterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and Fraser,C.M.		

TITLE	Direct Submission
JOURNAL	Submitted (31-Jan-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	
source	Location/Qualifiers
	1..10677
	/organism="Caulobacter crescentus CB15"
	/strain="CB15"
gene	/db_xref="taxon:190650"
	220..2433
	/gene="CC3043"
CDS	220..2433
	/gene="CC3043"
	/note="identified by match to TIGR protein family HMM TIGR00198"
	/codon_start=1
	/transl_table=11
	/product="catalase/peroxidase"
	/protein_id="AAK25005.1"
	/db_xref="GI:13424687"
	/translation="METPARNTPMADKVEDNIACPMGHGGRPANRDMWPQSLRLE GLNHAPRNPMEGFADYAEAFKSLDLDAVVDLHALMTDSQEWPAADFHYGGGLFIR LAWHAAGTVRTDGRGAGGQORFAPLNSWPDNTNLDKARLLWP IKQYGAKLWSA DLYLVNVALESMGFKTFAGGRADOWPEPELYWGPSTWLDKDRKSGERELDSPL GAVOMGLIYNDEGNGNDPLASARDIRETFARMANDEETVALIAGTFTGKAHGA GDASLVGPEGAIERQGFGNWASKTGKGDALITGGPEVITOTPTPEWSNHFENL FKEWELTQSPAGAKOWAKNAPADIPDAPFNKTHVPRMLTSDLRALRDPAYEKISR REFENPDQADAFARAFWKLTTHRDMPGPIGRYLGPIVPRKEELIWQDPIPAVDHPLADDK DIAALKAKLATGLSASDLVSTAWASASTYROSDKRGGANGARIRLAPQKDWANVNP VLAKVLALEGVQKDFNASAGGKKISLADLIVLAGAAAIKRAAKADAGTSVTPFPAG RMDASAEOTDAHSFEALPRSDGRNRYGPKHNAPEALVDRQAQLLGLSPELTVL WGLRVLGANADGSKDVTNRGALSNDFNLLSMETTWSPSTANAFAGHDKRSSE PRWTATRVLLIFGSHAEIRAPAEVYACADQSEKFCVDFVTANKVYNADRLDAA"
gene	complement(2685..4538)
CDS	/gene="CC3044"
	complement(2685..4538)
	/gene="CC3044"
	/note="identified by match to TIGR protein family HMM TIGR00110"
	/codon_start=1
	/transl_table=11
	/product="dihydroxy-acid dehydratase"
	/protein_id="AAK25006.1"
	/db_xref="GI:13424688"
	/translation="MPYRSRTTHGRNMGAGRLWRATCMKDFGKPIIAVANSTF QVPGHVHLKDLQVARELEAAGVAKENTIAVDGIAHGGMGLYSLPSRLDIAD SVETMVAHCADAIYICNSCDKITPCNLAARLNIPVYVSGGPMACKVTVKGIIR ALDLDAMVVAADDSYSEVEAIEAKAECTCGSCGSMFTANNCILTEALGLSLFNG GSVLATHADREALFKEAGRVVVDLCORWEQEDATALPRGIATRAAFENAMSLDIAMG GSTNTVHLIAAAHEGIDFSMADIDRLSRHVPCLSKVAPAKSDVHEDVHRAGGYMA ILGELERGLIDASQPTVHAPTMEALARDWIDRTNSQIAHEFFKAAPGKPTQVAFS QARWEELDRENGVIYRVEHPFSKDGGLAVLGNLAPGECIVKTAGVDESILTPRG TARVFEESQAAVSGLLGGOVKAGEVVVIRYEGPKGGPGQEMLYPTTYLKSGLGAAC ALVTDGRFSGTSGLSTGHVSPPAGGGLIALVETGDPILIDITPTGTLVSDAVLA ARREAQIARKGADWTPLNRKDLTPALRAAAMTTNAARGAVRDVDSQIERG"
gene	4637..5527
	/gene="CC3045"
CDS	4637..5527
	/gene="CC3045"
	/note="identified by match to protein family HMM"
	/codon_start=1
	/transl_table=11
	/product="isoleucine biosynthesis transcriptional activator"
	/protein_id="AAK25007.1"
	/db_xref="GI:13424689"
	/translation="MDIRQFHPFAVAETLHFGRARLGIOTPTLSQSLQALEKALG APLEFATKRHVLTALGRQMLPHVLEALAAVDALPDTARRLRDGTGTLSSVSTAD RYSLPDLVRYAEAFPGVEIQLEVTADVDQVPAIQAGERHAGIIPPNSSVPAALAY RRLVSEPLVAVTPEAAGEPLDIAALADAPLVLPFTVAPAFHDLVTGVVAARGQV RIVQEATQOMOTIILSVSAGLGMALAPASLRKARLVGRVYVDLVDPPILETLGVWRDE AAPTLQGLILLETGDSAPD"
gene	5578..6093
	/gene="CC3046"

QY	688	ATCCCTGGGGGGCTGATTTTGATATATGATCCACCAAGATTTCAACAGCTGGATATGAGGCTC	747
Db	167	ACCCGATGGGTGGGGCTTCTGACTATGTCGCGGAGGTCTGCGACCATCGAGCTTGACGCC	226
QY	748	TGAAAAAAGATATCAAGATTTCTGACAACTTCCAGGATTTGTCCTCGCGGATATG	807
Db	227	TGACGGGACATCGAGGAAGTATGACACCTCGACCGCTGTCGCGCGCGGACTACG	286
QY	808	GTCAATATGCTCTTCTTTATTCGATGCTTGGCAGGTCGCGGACATACAGGACAT	867
Db	287	GCCACTACGGCGCTTATTCGGATGCTGGCAGCTGCGGACCTACCGCATCC	346
QY	868	ATCATGGCGGGAGCGGCGGCTGGTGCAGCAAGCTTTTGACCGCTGAACAGCTGC	927
Db	347	ACACGGCGCGCGCGCGGCGGCGGATGACGAGCTTTCGCGCGCTTAACAGCTGC	406
QY	928	CGGATAACGTTAATCTGATAAAGCCGCTGATTTGCTGGCCAGTTCGTCGCGGATATG	987
Db	407	CCGACACCGCAGCTTGGACAGCGCGCGCTGCTGTCGCGGTCAGAGAGTACG	466
QY	988	GCTCCAGTATTTCTCGGGAGACCTGATGTCCTGACTGTGTAATGTTGCCCTTGAATCCA	1047
Db	467	GCAAGAAGCTCTCATGGCGGAGCTGATTTGCTGCGCGCAACTGCGCGCTGGAATCGA	526
QY	1048	TGGGATTTAAACGCTGGGATTTGCTGCGCGGAGAGAGATGACTGGGAGTCCGACCTGG	1107
Db	527	TGGGCTTCAAGACGCTGGGCTTGGCTTGGCGCGGTCGACCACTGGGAGCGCGGATGAG	586
QY	1108	TATACTGGGGGGCTGACAAAGCCCTTCTGACATAACCGGGATA--AAACCGGAAAC	1164
Db	587	TCATTTGGGCAAGGAGCCACCTGGCTGGCGGATGAGCTTACGGGTAAACGGGATC	646
QY	1165	TTGAGAACTCTTGGCGGACAGCTGATGCTCTGACTGATGTTGCTGCTTGAATCCA	1224
Db	647	TGAGAAACCGCTGGCGGCTGACAGTGGGCTGATCTACGTGAACCGGAGGCGCA	706
QY	1225	GTGAAACACGATCTCTGCTTGGCTTGGCGGAAATATATCAGGAGCTTTTTCAGTATG	1284
Db	707	ACGGCAACCGGAGCCCATGGCGCGGCTGACATTCGCGAGAGCTTTTCGCGCATGG	766
QY	1285	CCATGATGATGAGGAGCTGTCGCTGATCTGCGGAGGCGATACATTTGTTAAAGCAC	1344
Db	767	CCATGAGACGCTGGAACAGCGGCTGATCTGCGGCTGACACTTTCGGTAAAGCC	826
QY	1345	ATGGTGAGGCTCTCTGAAAATGATTTGGCGAGGCTGATGTCACCTGTGAGG	1404
Db	827	ATGGCGCG--GCCCGCGGATCTGCTGCGCGCCGAAACCGAGGCTGCTCCGCTGAGC	883
QY	1405	AGCAGGAGCTGGGATGAAAAATTAATGTTGTCAGAGAAACGCAATATACCATCACA	1464
Db	884	AGATGGGCTTGGGCTGGAAGAGCTGATGGCACCGGACCGGTAAGGAGCGCATCACA	943
QY	1465	GTGGCTTGAAGGAGCTTGTGCA 1488	
Db	944	CGGCGATCGAGTCTGATGACGA 967	
RESULT 33			
AF314115		2223 bp	DNA linear BCT 11-FEB-2001
LOCUS			
DEFINITION			Mycobacterium tuberculosis isolate F16 catalase-peroxidase (katG)
VERSION			gene, complete cds.
AF314115			
KEYWORDS			AF314115.1 GI:12744169
SOURCE			Mycobacterium tuberculosis.
ORGANISM			Mycobacterium tuberculosis.
			Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
			Actinomycetales; Corynebacterineae; Mycobacteriaceae;
			Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE			1. (bases 1 to 2223)
AUTHORS			Orru, G., Iona, E., Memmi, G., Oggioni, M.R., Fattorini, L., Orefici, G.
TITLE			and Pozzi, G.
JOURNAL			Mutation Associated to Isoniazid Resistance in Italian Isolates of
REFERENCE			Mycobacterium tuberculosis
AUTHORS			Unpublished
TITLE			2 (bases 1 to 2223)
JOURNAL			Orru, G. Submission
FEATURES			Direct Submission
source			Submitted (18-Oct-2000) Dip. Biologia Molecolare, Laboratorio di
			Microbiologia Molecolare e Biotecnologia, Università di Siena,
			Viale Bracci, Siena 53100, Italy
			Location/Qualifiers
			1..2223
			/organism="Mycobacterium tuberculosis"
			/isolate="F16"
			/db_xref="taxon:1773"
			/note="isoniazid resistant isolate"
gene			1..2223
			/gene="katG"
CDS			1..2223
			/gene="katG"
			/codon_start=1
			/transl_table=11
			/product="catalase-peroxidase"
			/protein_id="AAK06517.1"
			/db_xref="GI:12744170"
			/translation="MPEQHPITETTTTGAASNGCPVVGHNKYPVEGGNODWPNRLN
			LKYLHONPAVDPMGAADFAAEVATIDVALTRDIEVNTTSQWMPADYGHVPLF
			IRMAHAAATYRIHGRGAGGAMORFAPLNSWPDNASLDKARLLWPVKYKGGKLS
			WADLIYPAGCALESMEKFTFGFCGFRVDQWEDVYWGKEATWLDERSGRDLN
			PLAIVOMGLIYVNPENPGNPDMAAVIDIREFRMAMNDVETALIVGGHTFGKTH
			ILYGYEELTKSPAGQWTKADGAGTTPDFPGGSRSPKTLATDLSLRVPIYER
			ITRRWELPELAEFAKAWYKLIHRDMGVARYLGLVPKQTLWDPVPVYSHDLV
			GEAEIASLKSQIRASGLTVSOLVSTAWAASSFRGSKRGANGRIQLPOVGHVW
			DPGDLKRVLTLEIOESFNSAAGNIKVSFADLYVLGSCAIEAKAAKAAGNIYTP
			FTPTDASQBDQDVFVLEPKADCFRNLGKNPLPAEYMLDKANLILTSAPEM
			TVLVGLRVIGANTKRLPIGFTFEASESLTNDFVNLDDMGITWEPSPADDTGYGKD
			GSGKVKWTSRVLGFSNELSELVEYGADDAQPKFVQDFVAANDKVNMLDRFDYR
			"
BASE COUNT	427 a	695 c	742 g 359 t
ORIGIN			
Query Match		20.08;	Score 298.4; DB 1; Length 2223;
Best Local Similarity		62.48;	Pred. No. 1.8e-77;
Matches	502;	Conservative	0; Mismatches 296; Indels 6; Gaps 2;
QY	688	ATCCCTGGGGGGCTGATTTTGATATGATCCACCAAGATTTCAACAGCTGGATATGAGGCTC	747
Db	167	ACCCGATGGGTGGGGCTTCTGACTATGTCGCGGAGGTCTGCGACCATCGAGCTTGACGCC	226
QY	748	TGAAAAAAGATATCAAGATTTCTGACAACTTCCAGGATTTGTCCTCGCGGATATG	807
Db	227	TGACGGGACATCGAGGAAGTATGACACCTCGACCGCTGTCGCGCGCGGACTACG	286
QY	808	GTCAATATGCTCTTCTTTATTCGATGCTTGGCAGGTCGCGGACATACAGGACAT	867
Db	287	GCCACTACGGCGCGCTTATTCGGATGCTGGCAGCTGCGGACCTACCGCATCC	346
QY	868	ATCATGGCGGGAGCGGCGGCTGGTGCAGCAAGCTTTTGACCGCTGAACAGCTGC	927
Db	347	ACACGGCGCGCGCGCGGCGGCGGATGACGAGCTTTCGCGCGCTTAACAGCTGC	406
QY	928	CGGATAACGTTAATCTGATAAAGCCGCTGATTTGCTGGCCAGTTCGTCGCGGATATG	987
Db	407	CCGACACCGCAGCTTGGACAGCGCGCGCTGCTGTCGCGGTCAGAGAGTACG	466
QY	988	GCTCCAGTATTTCTCGGGAGACCTGATGTCCTGACTGTGTAATGTTGCCCTTGAATCCA	1047
Db	467	GCAAGAAGCTCTCATGGCGGAGCTGATTTGCTGCGCGCAACTGCGCGCTGGAATCGA	526
QY	1048	TGGGATTTAAACGCTGGGATTTGCTGCGCGGAGAGAGATGACTGGGAGTCCGACCTGG	1107

Db 527 TGGGCTTCAAGACGCTTCGGCTTCGGCTTCGGCCGGTCGACCATGGGAGGCCCGCATGAGG 586
 QY 1108 TATACCTGGGGGCTTGACAAACAGCCTCTTGACAGATAACCGGGATA---AAAACGGGAAAC 1164
 Db 587 TCTATTGGGGCAAGAACCCACCTGGCTTCGGCGATGAGCGTTACAGCGGTAAGCGGATC 646
 QY 1165 TTCAGAAACCTCTGCGCCACGACAGATGGGACTTATTATGTCATTCCTGAAGGCCCG 1224
 Db 647 TGGAGAACCCCTGCGCGCTGAGATGGGCTGATCTAGTGAACCCCGAGGGCCGA 706
 QY 1225 GTGAAAACACAGATCTCTGCTTCGCGGAAAGATATCAGGAAGCTTTTTCACGTATGG 1284
 Db 707 ACGCAACCCGACCCCATGCCCCGGGTGCGACATTCGCGAGACGTTTCGCGGCATGG 766
 QY 1285 CCATGGATGATGAGAGACTGTGCCCTGATGCGGGAGGCGATACATTTGGTAAAGCAC 1344
 Db 767 CATTGAAGAGCTGCGAAACAGCGCGCTGATGCTGCGCGGTTCACACTTCGCTGAAGACC 826
 QY 1345 ATGTGACAGCTCTCCTGMAAATGATTTGGCGCAGGCGCTGATGTGACCTGTGGAGG 1404
 Db 827 ATGGCGCG---GCCGCGCGATCTGGTCGCGCCCGAAACCCGAGGCTGCTCCGCTGGAGC 883
 QY 1405 AGCAGGACTGGATGGAATAAATGCTGTACAGGAACGGCAATATACCATCACCA 1464
 Db 884 AGATGGGCTTGGCTGGAAGAGCTGATGCGACCGAACCCTGAAGGACCGCATCACCA 943
 QY 1465 GTGGCTGGAAGAGCCTGGTCA 1488
 Db 944 CGGCATCGAGGTCGTATGGACGA 967

RESULT 34

AF314117

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

gene

CDS

AF314117 2223 bp DNA linear BCT 11-FEB-2001
 Mycobacterium tuberculosis isolate F15 truncated
 catalase-peroxidase (katG) gene, complete cds.

AF314117.1 GI:12744173
 Mycobacterium tuberculosis.

Mycobacterium tuberculosis.
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;
 Mycobacterium; Mycobacterium tuberculosis complex.

1 (bases 1 to 2223)
 Orru, G., Iona, E., Memmi, G., Oggioni, M. R., Fattorini, L., Orefici, G.
 and Pozzi, G.

Mutation Associated to Isoniazid Resistance in Italian Isolates of
 Mycobacterium tuberculosis

Unpublished
 2 (bases 1 to 2223)

Orru, G.
 Direct Submission

Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di
 Microbiologia Molecolare e Biotecnologia, Università di Siena,
 Viale Bracci, Siena 53100, Italy

Location/Qualifiers
 1..2223

/organism="Mycobacterium tuberculosis"
 /isolate="F15"
 /db_xref="taxon:1773"

/note="isoniazid resistant isolate"
 1..2223

/gene="katG"
 1..84

/gene="katG"
 /note="KatG"

/codon_start=1
 /trans_table=11

/product="truncated catalase-peroxidase"
 /protein_id="AAK06519.1"

/db_xref="GI:12744174"
 /translation="MPEQHPPTITETTTGAASNGCPVVGHMK"

BASE COUNT 429 a 694 c 741 g 359 t

ORIGIN

Query Match 20.0%; Score 298.4; DB 1; Length 2223;
 Best Local Similarity 62.4%; Pred. No. 1.8e-77;
 Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGGTGATTTTGTATGCCCACAGATTTTCAACAGCTGGATATGGAGGTC 747
 Db 167 ACCGATGGTGGGGTTCGACTATCCGCGGAGGTCCGACCATCGACTTGAAGCCC 226

QY 748 TGAATAAGATATCAAGATTTGCTGACAACCTTCCAGGATTTGGTCCCTGCGGATATG 807
 Db 227 TGACGCGGACATCGAGGAAGTATGCCACCTCGCAGCCGTGTGGCCCGCGACTAG 286

QY 808 GTCAATATGCTCTTCTTTTATTCGTATGCTTGGCAGCGTCCCGAATACACAGAT 867
 Db 287 GCACCTACCGGCGCTGTTTATTCGGATGGCTGGCAGCTCCGCGACCTACCGCATCC 346

QY 868 ATGATGGCGGGAGGCGCAGTGTGTCAGCAACGTTTGAACCGCTGAACAGCTGGC 927
 Db 347 ACGACGGCGCGCGCGCGCGCGCGCATCGAGGTTTCGCGCGCTTAACAGCTGGC 406

QY 928 CGGATAACGTTTATCTGGATAAAGCCCGCTCGATGCTGTGGCCAGTCAAGAAAAATACG 987
 Db 407 CCACAAACCCAGCTTGGACAAGCGCGCGCGCTGTGTGGCGGTCAAGAAGAGTACG 466

QY 988 GCTCCAGTATTTCTTGGGAGACATGATGCTCTGACTGGTAATTTGCCCTTGAATCCA 1047
 Db 467 GCAAGNAGCTCTCANTGGCGGACCTGATGTTTTCGCGGCANCTGCGCGTGGAAATCGA 526

QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAAGTACTGGAGTGGACCTGG 1107
 Db 527 TGGGCTTCAAGACGTTTCGGGTTTCGGCGCGGTGACCCAGTGGGAGCCCGATGAGG 586

QY 1108 TATCTGGGGGCTGACACAAAGCCTCTTCAGATAACCGGGATA---AAACGGGAAAC 1164
 Db 587 TCTATTGGGGCAAGAACCCCTGCTGGCGGATGAGCGTTTACAGCGGTAAAGCGGATC 646

QY 1165 TTCAGAAACCTCTTGC CGCCACGACAGATGGGACTTATTATGTCATCTCTGAAGCCCG 1224
 Db 647 TGGAGAACCCGCTGCGCGCGGTGACAGTGGGCTGATCTAGTGAAACCGGAGGGCGCA 706

QY 1225 GTGAAAACACAGATCTCTGCTTCGCGGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
 Db 707 ACGGCAACCCGACCCCATGCGCGCGGTGCGACATTCGCGAGACGTTTCGCGCATGG 766

QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATGTCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
 Db 767 CCATGAACGACTCGAAGACGCGCGCTGATGTCGCGGCTCACACTTTCGGTAAGACCC 826

QY 1345 ATGGTGCAGCGTCTCTTGAATAATGATTTGGCGAGGCGCTCATGCTGTGGAGG 1404
 Db 827 ATGGCGCG---GCCCGCGGATCTGTCGCGCCCGGACCCGAGGCTGCTCCGCTGGAGC 883

QY 1405 ACGAGGACTGGATGGAATAAATAATGTTGTAAGAAACGCAATATACCATCACCA 1464
 Db 884 AGATGGGCTTGGGCTGGAAGAGCTGATGCGACCGGAACCGTAAGACCGCATCACCA 943

QY 1465 GTGGCTGGAAGGAGCCTGGTCA 1488
 Db 944 GCGGCATCGAGTCTGTATGGACGA 967

RESULT 35

AF314120

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AF314120 2223 bp DNA linear BCT 11-FEB-2001
 Mycobacterium tuberculosis isolate An05 catalase-peroxidase (katG)
 gene, complete cds.

AF314120
 AF314120.1 GI:12744179

Mycobacterium tuberculosis.

ORGANISM	Mycobacterium tuberculosis Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex. 1 (bases 1 to 2223)
REFERENCE	Orru, G., Iona, E., Memmi, G., Oggioni, M.R., Fattorini, L., Orefici, G. and Pozzi, G.
AUTHORS	Mutation Associated to Isoniazid Resistance in Italian Isolates of Mycobacterium tuberculosis Unpublished 2 (bases 1 to 2223)
TITLE	Direct Submission
JOURNAL	Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di Microbiologia Molecolare e Biotecnologia, Università' di Siena, Viale Bracci, Siena 53100, Italy
REFERENCE	1. .2223
AUTHORS	Location/Qualifiers
TITLE	source
JOURNAL	/organism="Mycobacterium tuberculosis" /isolate="An05" /db_xref="taxon:1773" /note="isoniazid resistant isolate"
REFERENCE	1. .2223
AUTHORS	/gene="katG"
TITLE	/gene="katG"
JOURNAL	/note="katG"
REFERENCE	/codon_start=1 /transl_table=11 /product="catalase-peroxidase" /protein_id="AAK08522.1" /db_xref="GI:12744180"
AUTHORS	/translation="MPEQHPTTITTTGAASNGCPVVGHMKYPVEGGNQDWWPNRLN LKVLPONPAVDPMGAADYAAEVATIDVALTRIDIEVNTISOPWMPADYHGYPLE IRMAHAACTVRIHPDRGAGSGMORAPLNSWPDNASLDKARELLWPVKKYKKLS WADLVFAGNCALESNGKRTGFGFGRVDQWEPYVWKEATWLDGERSGRDLN PLAAVOMGLIYVPEPNPNPMDAAVDIRETERFAMNDVETAAALIVGGHTFGKTH GAGPADLVPEPEAPLEQMGKWSYGTGDKDITSGIEVVVNTPTKWNDSLE ILYGEWELTSPAGAWYTKADGAGAGTIPDPFGGPRSPMTLATDLSLRVDPYIER IRWLEHPEELADEFAKAWYKLIHRDMGPVARYLGPVLPKQTLWODPVPVSHVDV GRAETASLQVRSOLVSTAWAASAFRSGPKRGANGRIRLOPOVGHVW DPDGLRVIRLEIQSFNSAPGNKIVSFADLVYLGCAAEKAAKAAGHNITVP FPGRTDASQEQTDVESFAVLEPRADGFRNLYKGNPLPAEYMLLDKANLLTUSQEM TVYVGLVRLGANYRLPLGYTFEASESLTNDFFNLDIMGITWEPSPADDTGYQGD GSGKVKVTSRVDLVVGSNSRLALVEYVIGADDAQPKFVQDFVAADKVNMLDRFVDR "
BASE COUNT	428 a 695 c 742 g 358 t
ORIGIN	
Query Match	20.0%; Score 298.4; DB 1; Length 2223;
Best Local Similarity	62.4%; Pred. No. 1.8e-77;
Matches	Conservative 0; Mismatches 296; Indels 6; Gaps 2;
Qy	688 ATCCCTGGGGGCTGATTTGATATGACACAGATTTCAACAGCTGGATGAGGCTC 747
Db	167 ACCCGATGGCTGGCGGTTGACTACTATCGCGGAGTTCGCGACCATCGACGCTTACGCGCC 226
Qy	748 TGAATAAGATATCAAGATTTCTTCACAACTTCCAGGATTTGGTCCCTCGGATATG 807
Db	227 TGACGGCGGACATCGAGGAAGTATGACACCATCGCAGCCGCTGGTGGCCCGGACACTACG 286
Qy	808 GTCAATATGCTCTTCTTTATTCATGATGCTTGGCTGGCAGGCTGCGGGAACATACAGGACAT 867
Db	287 GCACATACGGCGCTGTTTATTCGGATGGCTGGCAGCTGCGGCGACCTACCGCATCC 346
Qy	868 ATGATGCGCGGGAGCGCCAGTGGTGTGTCAGCAACGTTTTGAACCGCTGAACAGCTGGC 927
Db	347 ACGACGCGCGCGCGCGCGCGGCGGCGGATCGACGCTGTCGCGCGCTTAACAGCTGGC 406
Qy	928 CGGATACGTTAATCTGGATTAAGCCCGCTGCGATTCCTGTGGCCAGTCAAGAAAATACG 987
Db	407 CCGACACACGAGCTTGGACAGGCGCGCGCTGCTGTGGCGGCTCAAGAAGATACG 466
Qy	988 GCTCCAGTATTTCTTGGGGAGACCTGATGGTCTGACTGATGTTGTCCTGGAATPCA 104
Db	467 GCAAGAAGCTCTCATGTGGCGGACCTGATTTGTTTCGCGGCAACTGCGCGCTGGAATPCA 526
Qy	1048 TGGGATTTAAACAGCTGGGATTTGCTGCGGAAAGAGATGACTGGAGTTCGGACCTGG 1107
Db	527 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCGGCTGACCACTGGGAGCCCGATGAGG 586
Qy	1108 TATACTGGGGGCTGACAAACAGCTCTTTCAGATAACCCGGGATA---AAAAAGGGAAC 1164
Db	587 TCTATTGGGCAAGAACCCACCTGGCTCGCGGATGAGCGTTACAGCGGTAACGGGATC 646
Qy	1165 TTCAGAAACCTTCTGCCGCCACGAGATGGGACTTATTTATGTCATTCCTGGAAGGCCCG 1224
Db	647 TGGAGAACCGCTGCGCGGCTGAGATGGGCTGATCTACGTGAACCCGAGGGGCGCA 706
Qy	1225 GTGGAAAACCATGCTCTGCTTCCGCGAAGATATCAGGAAGCTTTTTCAGTATGG 1284
Db	707 ACGGCAACCCGACCCCATGCGCGCGGCTGACATTCGCGGAGACGTTTCGGGCGCATGG 766
Qy	1285 CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGGAGGAGGCATACATTTGGTAAAGCAC 1344
Db	767 CCATGACGACGTCGAAACAGCGGCGCTGATCGTCGGGCTCACACTTCGTTAAGACCC 826
Qy	1345 ATGGTGCAGCGTCTCTGAAAATGATTTATGGCGGAGGCTGTGTTGTCACCTGTGGAGG 1404
Db	827 ATGGCGCGG---GCCCGCGCATCTGTGTCGCGCCGAAACCCGAGGCTGCTCGCGTGGAGC 883
Qy	1405 ACCAGGGACTGGATGGAAATAAATGCTGTACAGGAAGCGCAATATACCATCACCA 1464
Db	884 AGATGGGCTTGGCTGGAGAGCTGCTGATGGCACCGGAACCGGTAAGACGCGATCACCA 943
Qy	1465 GTGGCTGGAAGGAGCGCTGCTGA 1488
Db	944 GCGGCATCGAGTCTGATGGACGA 967
RESULT	36
AF314121	2223 bp DNA linear BCT 11-FEB-2001
LOCUS	Mycobacterium tuberculosis isolate Rm09 catalase-peroxidase (katG)
DEFINITION	gene, complete cds.
ACCESSION	AF314121
VERSION	AF314121.1 GI:12744181
KEYWORDS	Mycobacterium tuberculosis.
SOURCE	Mycobacterium tuberculosis.
ORGANISM	Mycobacterium tuberculosis. Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex. 1 (bases 1 to 2223) Orru, G., Iona, E., Memmi, G., Oggioni, M.R., Fattorini, L., Orefici, G. and Pozzi, G.
REFERENCE	Mutation Associated to Isoniazid Resistance in Italian Isolates of Mycobacterium tuberculosis Unpublished 2 (bases 1 to 2223)
AUTHORS	Orru, G.
TITLE	Direct Submission
JOURNAL	Submitted (18-OCT-2000) Dip. Biologia Molecolare, Laboratorio di Microbiologia Molecolare e Biotecnologia, Università' di Siena, Viale Bracci, Siena 53100, Italy
REFERENCE	1. .2223
AUTHORS	Location/Qualifiers
TITLE	source
JOURNAL	/organism="Mycobacterium tuberculosis" /isolate="Rm09" /db_xref="taxon:1773" /note="isoniazid resistant isolate"
REFERENCE	1. .2223
AUTHORS	/gene="katG"
TITLE	/gene="katG"
JOURNAL	/note="katG"
REFERENCE	/codon_start=1 /transl_table=11 /product="catalase-peroxidase" /protein_id="AAK08522.1" /db_xref="GI:12744180"
AUTHORS	/translation="MPEQHPTTITTTGAASNGCPVVGHMKYPVEGGNQDWWPNRLN LKVLPONPAVDPMGAADYAAEVATIDVALTRIDIEVNTISOPWMPADYHGYPLE IRMAHAACTVRIHPDRGAGSGMORAPLNSWPDNASLDKARELLWPVKKYKKLS WADLVFAGNCALESNGKRTGFGFGRVDQWEPYVWKEATWLDGERSGRDLN PLAAVOMGLIYVPEPNPNPMDAAVDIRETERFAMNDVETAAALIVGGHTFGKTH GAGPADLVPEPEAPLEQMGKWSYGTGDKDITSGIEVVVNTPTKWNDSLE ILYGEWELTSPAGAWYTKADGAGAGTIPDPFGGPRSPMTLATDLSLRVDPYIER IRWLEHPEELADEFAKAWYKLIHRDMGPVARYLGPVLPKQTLWODPVPVSHVDV GRAETASLQVRSOLVSTAWAASAFRSGPKRGANGRIRLOPOVGHVW DPDGLRVIRLEIQSFNSAPGNKIVSFADLVYLGCAAEKAAKAAGHNITVP FPGRTDASQEQTDVESFAVLEPRADGFRNLYKGNPLPAEYMLLDKANLLTUSQEM TVYVGLVRLGANYRLPLGYTFEASESLTNDFFNLDIMGITWEPSPADDTGYQGD GSGKVKVTSRVDLVVGSNSRLALVEYVIGADDAQPKFVQDFVAADKVNMLDRFVDR "
BASE COUNT	428 a 695 c 742 g 358 t
ORIGIN	
Query Match	20.0%; Score 298.4; DB 1; Length 2223;
Best Local Similarity	62.4%; Pred. No. 1.8e-77;
Matches	Conservative 0; Mismatches 296; Indels 6; Gaps 2;
Qy	688 ATCCCTGGGGGCTGATTTGATATGACACAGATTTCAACAGCTGGATGAGGCTC 747
Db	167 ACCCGATGGCTGGCGGTTG

```

/codon_start=1
/transl_table=11
/product="catalase-peroxidase"
/protein_id="AAK06523.1"
/db_xref="GI:12744182"
/translation="MPEQHPPIITETTGAASNGCPVVGHMKYPVVEGGNODWPNRLN
LKVLHNPVADPMGAADFAAEVATIDVDAITRDIEVMTTSQWMPADYGHYGLF
IRMAWHAAGTVRIHDGAGGAGGQRFAPLNSWPDNALSIDKARLLNPPVKYKGLS
WADLVFAGNCALESMTGFTGFGVRDQWEPDEYWGKEATWLDRIYVGGTGTGKH
PLAQQVOMGLIIVNPEGNPDMAAAVDIRETFRMAMNDVETALIVGGHTFGKTH
GAGPADLVGPEPEAPLEOMGLCHKSSYGTGTGKDAITSGIEVWVWNTPTKWNSELE
ILYGYHELTKSPAGANQYATKDGAGAGTIPDFGPGRSPTMLATDLSURVDPIYER
ITRWLHEPELADFAKAWYKLIHRDMGPVARYLGLPKQTLMLWDPVPAYSHDLV
GEATIASLKSOIRASGLTVSOLVSTAWAASSFRGSDKRGAGRIRLQPVQWVNV
DEGDLRKVIRILEIESFNAPGNIKVSFADLVVLGGCAAIKAAKAAGHNITVP
PTPGRDASQOTDVEFPAPKPRGPNYLVGKGNPLPAEYMLDKANLLTSLAPEM
TVLVGGLRVLGANKRLPLGVFTSESLTNDFFVLLDMGITWEPSPADGTQGGK
GSGVKVKTGSRVDLVFVSGNSLRALVEYVGDADDAQPKFVQDFAAWDKVMNLDKDFVR
"
BASE COUNT      429 a   695 c   740 g   359 t
ORIGIN
Query Match      20.0%; Score 298.4; DB 1; Length 2223;
Best Local Similarity 62.4%; Pred. No. 1.8e-77;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
QY 688 ATCCCTGGGGGCGCTGATTTGATTCACACAGATTTCACACGCTGGATATGAGGCTC 747
DB 167 ACCCGATGGTGGCGGCTTGACTATGCCGGAGGTGCGACCATCGACGTTGACGCC 226
QY 748 TGAARAAGATATCAAGATTGCTGACACTCCAGGATTTGGTCCCTGCGGATATG 807
DB 227 TGACGCGGACATCAGGAAGTAGATCACCTCGACGCGTGGTGGCCCGCGCACTACG 286
QY 808 GTCATTATGTCCTTCTTTATGCTATGCTTGGCACGCTGCGGAAACATACAGACAT 867
DB 287 GCCACTAGGCGCGCTGTTATCGGATGGTGGACGCTGCGGACCATCGGATCC 346
QY 868 ATGATGCGCGGGAGGCGCGAGTGGTGCAGAACCTTTTGAACCGCTGAACAGCTGGC 927
DB 347 ACGACGCGCGGCGCGCGCGGCGGCATGACAGCGTTTCGCGCGCTTAACAGCTGGC 406
QY 928 CGGATAAGCTTAATCTGATTAAGCCGCTGATGCTGTGGCCAGTCAAGAAAATACG 987
DB 407 CCGAACAGCGCGCTTGACAAAGCGCGCGGCTGCTGTGGCGCGTCAAGAAAGTACG 466
QY 988 GCTCCAGTATTCTCGGGAGACCTGATGCTGCTGCTGATGCTGCTGCTGCTGCTGCTG 1047
DB 467 GCAAGAAGCTCTCATGCGGCGACTGATGTTTTCGCGCGCACTGCGCGCTGATCGA 526
QY 1048 TGGGATTTAAACCGCTGGATTTGCTGCGCGAAGAGATGACTGGGAGTGGGAGCTGG 1107
DB 527 TGGGCTTCAAGACGCTTCGGGTTGCGGTTTCGCGCGGTCGACACAGTGGGAGCGCG 586
QY 1108 TATACTGGGCGCTGACAAACGCTCTTCAGATTAACCGGATA---AAACCGGAAC 1164
DB 587 TCTATTGGGCAAGGAAGCAACCTGCGCTCGGCGATGAGCGTTACACGCGTAAGCGG 646
QY 1165 TTCAGAAACCTTTTCGCGCGCACGATGGGACTTATTATGTCAATCTCTGAAGCGCGCG 1224
DB 647 TGGAGAACCCTGCGCGCGGTGCAGATGGGCTGATCTACGTGAACCGCGAGGCGCGCA 706
QY 1225 GTGAAAACAGATCTCTGCTGCTGCTGCGCGAAGATATCAGGGAAGCTTTTTCACG 1284
DB 707 ACGCAACCGGACCCCATGGCGCGCGGCTGCACATTCGCGAGAGCTTTGCGCGCATGG 766
QY 1285 CCATGGATGATGAGGAGACTGTGCGCTGATCCGGGAGGCGATACATTTGTTAAGACAC 1344
DB 767 CCATGAACAGCTCGAAACAGCGCGCTGATCTCGCGCGTCAACATTTGCTGTAAGACC 826
QY 1345 ATGTCGAGCGCTCTCTGTAAGAAATGATTGGCGCAGGCGCTGATGTCGACCTGTGCGAG 1404
DB 827 ATGCGCGC---GCCGCGCGGATCTGTGTCGCGCGCGGACCCGAGCGGCTGCTCGCGT 883

```

1405 AGCAGGACTGGGATGAAAAATAATGTGTACAGAAACGCAATATACATCACCA 1464
 884 AGATGGCTGGCTGGAGAGCTCGTATGGCAGCCGGAACCGGTAAGGACGCGATCACCA 943
 1465 GTGGCTGGAAGGAGCTGGTGA 1488
 944 GCGCATCGAGGCTGATGACGA 967

MTU40593 2223 bp DNA linear BCT 07-JAN-1996
 Mycobacterium tuberculosis INH-resistant strain 15726/89, mutant
 catalase-peroxidase (katG) gene, complete cds.

U40593
 U40593.1 GI:1150695

Mycobacterium tuberculosis strain=15726/89.
 Mycobacterium tuberculosis
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;
 Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE
 AUTHORS Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.
 TITLE katG gene mutations in Isoniazid-resistant Mycobacterium
 tuberculosis isolated from Finnish patients
 JOURNAL Unpublished
 2 (bases 1 to 2223)
 AUTHORS Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-1995) Harri J. Marttila, Mycobacterial Reference
 Laboratory, National Public Health Institute, Kivimyllynkatu 13,
 Turku, SF-20520, Finland

FEATURES
 Location/Qualifiers
 1..2223
 /organism="Mycobacterium tuberculosis"
 /strain="15726/89"
 /db_xref="taxon:1773"
 /note="INH-resistant strain"
 1..2223
 /gene="katG"
 /note="mutant katG"
 /codon_start=1
 /transl_table=11
 /product="catalase-peroxidase"
 /protein_id="AAK06523.1"
 /db_xref="GI:1150695"

gene
 1..2223
 /gene="katG"
 /note="mutant katG"
 /codon_start=1
 /transl_table=11
 /product="catalase-peroxidase"
 /protein_id="AAK06523.1"
 /db_xref="GI:1150695"

BASE COUNT 428 a 696 c 741 g 358 t
 ORIGIN

Query Match 20.0%; Score 298.4; DB 1; Length 2223;
 Best Local Similarity 62.4%; Pred. No. 1.8e-77;
 Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
 QY 688 ATCCCTGGGGGCGCTGATTTGATTCACACAGATTTCACACGCTGGATATGAGGCTC 747

Db 587 TCTATTGGGCAAGAAAGCCACCTGGCTCGCGCATGAGCGTTACAGCGTAAACGGGATC 646

QY 1165 TTCAGAAACCTCTTCCGCCAGCAGATGGGACTTATTATGTCAATCTCTGAAGGCCCG 1224

Db 647 TGGAGAACCCCTGGCCGCGTGCAGATGGGCTGATCTACGTGAACCCGGAGGCCGA 706

QY 1225 GTGGAACACGATCTCTGCTTCCGCGAAGATATCAGGAAGCTTTTTCACGTATGG 1284

Db 707 ACGCAACCCGACCCATGCGCGCGCTGCACATTCGAGAGCTTTTCGCGCATGG 766

QY 1285 CCATGGATGAGGAGACTGTGCCCTGATCGCGGAGGCATACATTTGGTAAAGCAC 1344

Db 767 CCATGAGGAGCTGGAACAGCGCGCTGATCGCGCGTCACTTTTCGGTAAGACCC 826

QY 1345 ATGTGCGAGGCTCTCTCTGAAATATGTTTGGCGAGGCGCTGATGTGCACTTGTGGAGG 1404

Db 827 ATGGCGCG--GCCCGCGCATCTGTGTCGCGCCGAAACCCGAGGCTGTCTCCGCTGGAGC 883

QY 1405 AGCAGGAGCTGGATGGAATAATATGTGTGACAGAAACGGCAATATACCATCACCA 1464

Db 884 AGATGGGCTTGGGCTGGAAGAGCTGATGTGCGACCGAACCCTGAAGGAGCGCATACCA 943

QY 1465 GTGGCTTGAAGAGGAGCTGTGTCGA 1488

Db 944 CGGCATCGAGTCTGTATGACGA 967

RESULT 39

MTU41306

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

RESULTS

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

RESULTS

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT 428 a 695 c 740 g 360 t

ORIGIN

Query Match 20.0%; Score 298.4; DB 1; Length 2223;

Best Local Similarity 62.4%; Pred. No. 1.8e-77;

Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGCTGATTTTGTATGTCACACAGATTTTCAACAGCTGGATGGAGGCTC 747

Db 167 ACCCGATGGGTGGGGTTCGACTATCGCGGAGTCCGACCATCGACGTTGACGCC 226

QY 748 TGAATAAGATATCAAGATTTGCTGACAACTTCCAGAGATTGTCCTCGGGATTATG 807

Db 227 TGACGCGGACATCGAGGAAGTATGACCACTTCGACCGTGTGGCCGCGCTACG 286

QY 808 GTCAATATGCTCTTCTTTTATGCTTATGCTTGCACGCTTCCGGAACATACAGACAT 867

Db 287 GCACATACGGCGCTGTTTATCCGATGGCTGCGACGCTCCGCGACCTACCCGATCC 346

QY 868 ATGATGGCGGGAGGCGCCAGTGTGTCAGAACGTTTGAACCGCTGAACAGCTGCG 927

Db 347 ACGACGCGCGCGCGCGCGCGCGCGCATCGACGCTTCCGCGCTTAAACAGCTGCG 406

QY 928 CCGATACGCTTAACTTGGATAAGCCCTCGATTGCTGTGGCCAGTCAAGAAAAATACG 987

Db 407 CCGACAAACCCAGCTTGGACAAAGCGCGCGCTGCTGTGGCGCGTCAAGAAAGATACG 466

QY 988 GCTCAGATATTTCTTGGGAGACCTGATGCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1047

Db 467 GCAAGAGCTCTCATGGCGGACCTGATGTTTTCGCGCACTTCCGCGCTGGATCGA 526

QY 1048 TGGGATTTAAACGCTGGGATTTGTCGGGGAAGAAAGATGACTGGGAGTGGACCTGG 1107

Db 527 TGGGCTTCAAGACGTTCCGGTTCCGGTTCGGCGGCTCGACGCTGGGAGCCGATGAGG 586

QY 1108 TATACCTGGGCGCTGACAAAGCCCTCTGTCAGATAACCGGATA---AAAACGGGAAC 1164

Db 587 TCTATTGGGGCAAGGAAGCCACCTGCTGCGGATGAGCGTTACAGCGGTAAAGCGGATC 646

QY 1165 TTCAGAAACCTCTGCGCCAGCAGATGGGACTTATTATGCTCAATCTGTAAGCCCG 1224

Db 647 TGGAGAACCCCTGCGCGCGTGCAGATGGGCTGATCTACGTGAACCCGGAGGCCGA 706

QY 1225 GTGGAACACGATCCTCTGCTTCCGGAAGATATCAGGAAGCTTTTTCACGTATGG 1284

Db 707 ACGGCAACCCGACCCCATGCGCGCGGCTGCACATTCGCGAGAGCTTTCGCGCATGG 766

QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGATACATTTGCTAAAGCAC 1344

Db 767 CCATGAAACGCTGCAACAGCGCGCTGATCTGCGCGCTCACACTTTCGTAAGACCC 826

QY 1345 ATGTGTCAGCGCTCTCTGAAATATGTTTGGCGAGGCGCTGATGTCACCTGTGGAGG 1404

Db 827 ATGGCGCG--GCCCGCGCATCTGTGCGCGCCGCAACCCGAGGCTGTCTCCGCTGAGT 883

QY 1405 ACGAGGAGCTGGATGGAATAATATGTTGTCAGGAACCGCAATATACCATCACCA 1464

Db 884 AGATGGGCTTGGGCTGGAAGAGCTCGTATGGCAACCGAACCCTGAAGGAGCGCATACCA 943

QY 1465 GTGGCTTGAAGGAGGCTGTGTCGA 1488

Db 944 CGGCATCGAGTCTGTATGACGA 967

RESULT 40

MTU41306

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

MTU41306 2223 bp DNA linear BCT 07-JAN-1996

Mycobacterium tuberculosis INH-resistant strain H0055/92,

catalase-peroxidase (katG) gene, complete cds.

U41306

U41306.1 GI:1150705

Mycobacterium tuberculosis strain-H0055/92.

/protein_id="AA85176.1"
/db_xref="GI:1150714"
translation="MPQHPPIITTTTGAASNGCPVVGHMKYPEVGGNODWPNRLN
LKVLPQAPVADPMPGAADYAAEVATIDVDALTRDIEVMTTSQWPAHYGHGGLF
IRMAHAAGTYRIHDGAGGAGGMOFAPLNSWPNASLDKARLLWPKKYGKKLS
WADLIVFAGNCALESMEGFTFGFGRVDQWEPDEVYWGKEATWLGDERYSKGRDLEN
PLAAYOMGLIYVNPENGPNDPMAAAVDIREFRMANNDVETALIVGGHFGKTH
GAGPADLVGPEPEAAAPLEQMGWKSQTDKDAITSGIEVWNTPTKWNDSFLE
ILYGEWELTKSPAGAWOYTKADGAGAGTTPDFGGRSPMTLADLSLRVDPIYER
ITRWLHPEELADEFAKAWYKLIHRDGMFVARYLGLVPKQTLWQDPVPVASHLV
DEADISLKQISQASGTSQVSTVAWAAASFGSKRGANGRIRLOPQVGVN
GPDGLKRVLTLEIEQESNSAAGNPKVFSADLVVLGCAALEKAAKAAGHNIVP
FTGRTDASQEDNDFSAVLEPKADGRNLYLKGPNLPAYMLLKANLLILSAPEM
TVLGGTLRGVGNKRLPLGFTFEASESLINDFFVNLDMGITWEPSPADGTYGKD
GSGKVKVTSGRVDLVFQPNSELRLVEVYGADDAQPKFQVDFVAANDKVMNLDREDFVR
"

BASE COUNT	428 a	696 c	741 g	358 t
ORIGIN				
Query Match	20.0%; Score 298.4; DB 1; Length 2223;			
Best Local Similarity	62.4%; Pred. No. 1.8e-77;			
Matches 502; Conservative	0; Mismatches 296; Indels 6; Gaps 2;			
QY	688	ATCCCTGGGGGCTGATTTGATTATCCGACAGATTTCACAGCTGGATATGAGGCTC	747	
DB	167	ACCGATGGGTGCGCGCTTCGACTATCGCGGAGGTCCGACCATCGAGCTTGACGCC	226	
QY	748	TGAAAAAGATATCAAGATTGCTGACAACTTCCAGAGTATGGTCCCTCGCGATTATG	807	
DB	227	TGACGGGGACATCGAGAAAGTATGACCACTCGCAGCGCTGGTGGCCCGCGACTACG	286	
QY	808	GTCATTATGTCCTTCTTATTCGTATGCTTGGCAGCGTGGCGGACATACAGACAT	867	
DB	287	GCCACTAGGGCGCGCTTTATCGGATGGGTGGACGCTCGCGCACCTACCGCATCC	346	
QY	868	ATGATGCGCGGGAGCGCCAGTGGTGGTCCAGCAAGCTTTTGAACCGCTGAACAGCTGC	927	
DB	347	ACGACGCGCGCGCGCGCGCGCGCGCATGCGCGGTTTCGCGCGCTTACAGCTGGC	406	
QY	928	CGGATACGTTATCTGGATAAGCCCGTGCATGCTGTGGCCAGTCAAGAAAAATACG	987	
DB	407	CCGACAACGCCAGCTTGGCAAGCGCGCGCGCTGTGTGCGCGGTCAAGAAAGTACG	466	
QY	988	GCTCCAGTATTTCTGGGAGACCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1047	
DB	467	GCAAGAGCTCTCATGGCGGACCTGATTTTTCGCGGCACTCGCGCTGGAATCGA	526	
QY	1048	TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAAGATGACTGGAGTCCGACCTGG	1107	
DB	527	TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGCGCGGTTCGACCACTGGGAGCCGATGAGG	586	
QY	1108	TATAGTGGGCGCTGACAAACGCTCTTGCAGATAACCGGATA---AAAACGGAAAC	1164	
DB	587	TCTATTGGGCAAGAAACCACTGGCTCGCGCATGAGCGTTACGCGGTAAACGGGATC	646	
QY	1165	TTCAAGAACCTTCGCGCCACAGATGGACTTATTATGTCATCTCTGAAGGCCCGG	1224	
DB	647	TGGAGAACCGCTGCGCGCGGTGCAGATGGGCTGATCTACGTTGAACCGCGGAGCCGA	706	
QY	1225	GTGAAACACAGATCTCTGCTGCTCGCGGAAGAAATATCAGGAAGCTTTTTCACGTATGG	1284	
DB	707	ACGCAACCGGACCCCATGCGCGCGGTCTGACATTCGCGAGAGCTTTCGCGCGATGG	766	
QY	1285	CCATGGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGCGCATACATTGTTAAGCAC	1344	
DB	767	CCATGAACGAGCTCGAAACAGCGCGCTGATCTCGCGGCTCACACTTCGTTAAGACCC	826	
QY	1345	ATGCTGACGCTCTCTGTAATAATGATTGGCCAGGCGCTGATGTGTCACCTGTGGAG	1404	
DB	827	ATGCGCGG---GCGCGCGGATCTGGTGGCGCGCGCGAACCAGGCTGCTCGCGTGGC	883	
QY	1405	AGCAGGACTGGGATGAAATAAATGTGGTACAGGAAACGCAATATACCATCACCA	1464	

Db	884	AGATGGCTGGCTGGAGAGCTCTGATGGCACCGGAACCGTAAAGCAGCGCATACCA	943	
QY	1465	GTGCGCTGGAAGAGCTCTGCTGA	1488	
DB	944	GCGCATCGAGCTCGTATGACGA	967	
RESULT 42				
MTU41313	2223 bp DNA linear BCT 07-JAN-1996			
LOCUS	Mycobacterium tuberculosis INH-resistant strain H0211/94,			
DEFINITION	catalase-peroxidase (katG) gene, complete cds.			
ACCESSION	U41313			
VERSION	U41313.1			
KEYWORDS	GI:1150719			
SOURCE	Mycobacterium tuberculosis strain-H0211/94.			
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
	Actinomycetales; Corynebacterineae; Mycobacteriaceae;			
	Mycobacterium; Mycobacterium tuberculosis complex.			
REFERENCE	1 (bases 1 to 2223)			
AUTHORS	Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.			
TITLE	katG gene mutations in isoniazid-resistant Mycobacterium			
JOURNAL	tuberculosis strains isolated from Finnish patients			
REFERENCE	2 (bases 1 to 2223)			
AUTHORS	Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-NOV-1995) Harri J. Marttila, Mycobacterial Reference			
	Laboratory, National Public Health Institute, Klimamyllynkatu 13,			
	Turku, SF-20520, Finland			
FEATURES	Location/Qualifiers			
source	1..2223			
	/organism="Mycobacterium tuberculosis"			
	/strain="H0211/94"			
	/db_xref="taxon:1773"			
gene	/note="INH-resistant strain"			
	1..2223			
CDS	/gene="katG"			
	1..2223			
	/gene="katG"			
	1..2223			
	/codon_start=1			
	/transl_table=11			
	/product="catalase-peroxidase"			
	/protein_id="AA85179.1"			
	/db_xref="GI:1150720"			
	/translation="MPEQHPPIITTTTGAASNGCPVVGHMKYPEVGGNODWPNRLN LKVLPQAPVADPMPGAADYAAEVATIDVDALTRDIEVMTTSQWPAHYGHGGLF IRMAHAAGTYRIHDGAGGAGGMOFAPLNSWPNASLDKARLLWPKKYGKKLS WADLIVFAGNCALESMEGFTFGFGRVDQWEPDEVYWGKEATWLGDERYSKGRDLEN PLAAYOMGLIYVNPENGPNDPMAAAVDIREFRMANNDVETALIVGGHFGKTH GAGPADLVGPEPEAAAPLEQMGWKSQTDKDAITSGIEVWNTPTKWNDSFLE ILYGEWELTKSPAGAWOYTKADGAGCTIPDPFGGRSPMTLADLSLRVDPIYER ITRWLHPEELADEFAKAWYKLIHRDGMFVARYLGLVPKQTLWQDPVPVASHLV GAEALSKQSIRASGLVSVSTAWAASFGSKRGANGRIRLOPQVGVN DPDGLKRVLTLEIEQESNSAAGNPKVFSADLVVLGCAALEKAAKAAGHNITVP FTGRTDASQEDNDFSAVLEPKADGRNLYLKGPNLPAYMLLKANLLILSAPEM TVLGGTLRGVGNKRLPLGFTFEASESLINDFFVNLDMGITWEPSPADGTYGKD GSGKVKVTSGRVDLVFQPNSELRLVEVYGADDAQPKFQVDFVAANDKVMNLDREDFVR "			
BASE COUNT	428 a	695 c	741 g	359 t
ORIGIN				
Query Match	20.0%; Score 298.4; DB 1; Length 2223;			
Best Local Similarity	62.4%; Pred. No. 1.8e-77;			
Matches 502; Conservative	0; Mismatches 296; Indels 6; Gaps 2;			
QY	688	ATCCCTGGGGGCTGATTTGATTATCCGACAGATTTCACAGCTGGATATGAGGCTC	747	
DB	167	ACCGATGGGTGCGCGCTTCGACTATCGCGGAGGTCCGACCATCGAGCTTGACGCC	226	
QY	748	TGAAAAAGATATCAAGATTGCTGACAACTTCCAGAGTATGGTCCCTCGCGATTATG	807	
DB	227	TGACGGGGACATCGAGAAAGTATGACCACTCGCAGCGCTGGTGGCCCGCGACTACG	286	
QY	808	GTCATTATGTCCTTCTTATTCGTATGCTTGGCAGCGTGGCGGACATACAGACAT	867	
DB	287	GCCACTAGGGCGCGCTTTATCGGATGGGTGGACGCTCGCGCACCTACCGCATCC	346	
QY	868	ATGATGCGCGGGAGCGCCAGTGGTGGTCCAGCAAGCTTTTGAACCGCTGAACAGCTGC	927	
DB	347	ACGACGCGCGCGCGCGCGCGCGCATGCGCGGTTTCGCGCGCTTACAGCTGGC	406	
QY	928	CGGATACGTTATCTGGATAAGCCCGTGCATGCTGTGGCCAGTCAAGAAAAATACG	987	
DB	407	CCGACAACGCCAGCTTGGCAAGCGCGCGCGCTGTGTGCGCGGTCAAGAAAGTACG	466	
QY	988	GCTCCAGTATTTCTGGGAGACCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1047	
DB	467	GCAAGAGCTCTCATGGCGGACCTGATTTTTCGCGGCACTCGCGCTGGAATCGA	526	
QY	1048	TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAAGATGACTGGAGTCCGACCTGG	1107	
DB	527	TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGCGCGGTTCGACCACTGGGAGCCGATGAGG	586	
QY	1108	TATAGTGGGCGCTGACAAACGCTCTTGCAGATAACCGGATA---AAAACGGAAAC	1164	
DB	587	TCTATTGGGCAAGAAACCACTGGCTCGCGCATGAGCGTTACGCGGTAAACGGGATC	646	
QY	1165	TTCAAGAACCTTCGCGCCACAGATGGACTTATTATGTCATCTCTGAAGGCCCGG	1224	
DB	647	TGGAGAACCGCTGCGCGCGGTGCAGATGGGCTGATCTACGTTGAACCGCGGAGCCGA	706	
QY	1225	GTGAAACACAGATCTCTGCTGCTCGCGGAAGAAATATCAGGAAGCTTTTTCACGTATGG	1284	
DB	707	ACGCAACCGGACCCCATGCGCGCGGTCTGACATTCGCGAGAGCTTTCGCGCGATGG	766	
QY	1285	CCATGGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGCGCATACATTGTTAAGCAC	1344	
DB	767	CCATGAACGAGCTCGAAACAGCGCGCTGATCTCGCGGCTCACACTTCGTTAAGACCC	826	
QY	1345	ATGCTGACGCTCTCTGTAATAATGATTGGCCAGGCGCTGATGTGTCACCTGTGGAG	1404	
DB	827	ATGCGCGG---GCGCGCGGATCTGGTGGCGCGCGCGAACCAGGCTGCTCGCGTGGC	883	
QY	1405	AGCAGGACTGGGATGAAATAAATGTGGTACAGGAAACGCAATATACCATCACCA	1464	

Db	227	TGACGGCGGACATCGAGGAAGTGATGACACCTCGCAGCGCTGGTGGCCCGCGACTACG	286
Qy	808	GTCAATATGTCCTTTCTTTATTCGTATGCTTGGCAGCTGGCGGAACATACAGGACAT	867
Db	287	GCCACTACGGCGCGCTTTATCCGATGCGTGGCAGCTGCGCGACCTACCGCATCC	346
Qy	868	ATGATGCGCGGAGCGGCAGTGGTGGTCAAGACGTTTTTGAACCGCTGAACAGCTGGC	927
Db	347	ACGACG	406
Qy	928	CGGATACGCTTAATCTGGATAAAGCCCGCTGATTTGCTGGCCAGTCAAGAAAAATACG	987
Db	407	CCGACACGCGCAGCTTGGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG	466
Qy	988	GCTCCAGTATTTCTTGGGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1047
Db	467	GCAAGAGCTCTCATGGCGGAGCTGATTTTCGCGCGCAACTGCGCGCTGGAATCGA	526
Qy	1048	TGGGATTTAAAGCTGGGATTTTCGCGGAGAGAGAGATGACTGGGAGTTCGGACCTGG	1107
Db	527	TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGCGCGGCTCGACCATGGGAGCCGATGAG	586
Qy	1108	TATACTGGGGGCTTGACAAAGCCTCTTCAGATACCGGGATA---AAACGGGAAC	1164
Db	587	TCTATTGGGCAAGGAAGCCACTGGCTCGCGCGATGAGCGTTACAGCGGTAAGCGGGATC	646
Qy	1165	TTCAAGAACCTCTTGCGCCACGACGATGGGACTTATTTATGTCATCTCGAAGGCCCGC	1224
Db	647	TGGAGAACCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCGGAGGGCCGA	706
Qy	1225	GTGAAACACGATCTCTGCTTCCGCGGAAAGATATCAGGAAGCTTTTTCACGTATGG	1284
Db	707	ACGCAACCGGACCCCATCGCGCGCGGTGCACATTCGCGAGACGTTTCGCGCGCATGG	766
Qy	1285	CCATGGATGATGAGGAGACTGTGCCCTGATCGCGGAGGCGATACATTTGGTAAAGCAC	1344
Db	767	CCATGACGAGCTGCAAGACGCGCGCTGATGTCGCGGTACACATTCGCGTAAAGACC	826
Qy	1345	ATGGTCAGCGCTCTCTGAAAAATGTATTGGCGCAGGGCTGATGTGTCACCTGTGGAGG	1404
Db	827	ATGGCGCG--GCCGCGCGATCTGTGTGCGCCCGCAACCCGAGGCTGTCTCCCTGGAGC	883
Qy	1405	ACGAGGACTGGGATGAAAAATAATGTGGTACAGAAACGGCAATAATACATCACCA	1464
Db	884	AGATGGCTTGGGCTGGAAGAGCTCGTATGGCAGCGAACCAGGTAAGGACGCGATACCA	943
Qy	1465	GTGCGCTGGAAGGAGCTCGTCA 1488	
Db	944	CGCGCATCGAGGTCGTATGACGA 967	
RESULT 43			
MTU41314			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Mycobacterium tuberculosis strain=H0725/93.			
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
Actinomycetales; Corynebacterineae; Mycobacteriaceae;			
Mycobacterium; Mycobacterium tuberculosis complex.			
1 (bases 1 to 2223)			
Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.			
katG gene mutations in isoniazid-resistant Mycobacterium			
tuberculosis strains isolated from Finnish patients			
Unpublished			
2 (bases 1 to 2223)			
Marttila,H.J., Soini,H., Huovinen,P. and Viljanen,M.K.			
Direct Submission			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

Laboratory, National Public Health Institute, Kiinamyllykatu 13, Turku, SF-20520, Finland			
Location/Qualifiers			
1. .2223			
/organism="Mycobacterium tuberculosis"			
/strain="H0725/93"			
/db_xref="taxon:1773"			
/note="INH-resistant strain"			
1. .2223			
/gene="katG"			
1. .2223			
/gene="katG"			
/note="mutant katG"			
/codon_start=1			
/transl_table=11			
/product="catalase-peroxidase"			
/protein_id="AA085180.1"			
/db_xref="GI:1150722"			
/translation="MPEQHPPIETTTGAASNGCPVYVGHKYPVEGGNQDWMNRLN LKVLHONPVAADPMGAADFAAEVATIDVDALTRDIEEVMTTSPQMPADYGHYGLF IRMAHAAGTYRIHDGKGAGGQRFAPLNSWPDNASLDPKARLLWPKVKYKKLS WADLVFAGNCALESMTGFTGFGFGRVDQWEPDEVYWGKEATWLGDERYSGRDLN PLAIVOMGLIYVNPENPNDPMAAADIRETFRMANNDVETALIVGSHHTFGKTH GAGPADLVPEEAPLEQMLGKMSYCTGCTKDAITSGIEVVTWPTTKWNDSFLE ILYGEWELKSPAGAWQYTAQDAGAGTIPDFGPGSRSPMLATDLSLRVDPIER ITRWLEHPEELADEFAKAWYKLIHRDMGPVARYLGPLVPKQTLMLMODPVPVSHDLV GEAEIASLSQIRASGLTVSOLVSTAWAASFRGSDKRGANGRRIRLOPQVWGYN DPQDLRKVIRTELEIQESNSAAGNPKVYFADLVVLGGCAIEAKAAAGHNITVP TFGRDASQEOBTVESFAVLEKPGDFRLYLGKGNPLPAEYMLLDKRAKMLTISAPEM TVLVGLRVLGANYKRLPLGVFTFEASESLNDFVNLDMGITWEPSPADGTYGKRD GSKVKTWTSRVDLVFGNSSELRALVEYTGADDAQPKFVDFVAAMDVKMNLDRFDR "			
BASE COUNT			
ORIGIN			
Query Match 20.08; Score 298.4; DB 1; Length 2223;			
Best Local Similarity 62.48; Pred. No. 1.8e-77;			
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;			
Qy	688	ATCCCTGGGGGCTGATTTTGATTTATGCCACAGATTTCACAGCTGGATATGGAGCTC	747
Db	167	ACCGATGGTGGCGGCTTCGACTATGCCGGGAGTCCGCGACCATCGAGCTTGACGCC	226
Qy	748	TGAATAAGATATCAAGATTGCTGACAACTTCCAGGATTTGGTCCCTCGCGGATATG	807
Db	227	TGACGGGGACATCGAGGAAGTATGATGCCACCTCGCAGCGGTGGTGGCGCGACTAG	286
Qy	808	GTCAATATGCTCTTTCTTTATTCGTATGCTGGCAGCTGGCGGAAACATACAGACAT	867
Db	287	GCCACTACGGCGCGCTGTTTATCCGGATGGCGTGGCAGCGTGC CGCGCACTCGCGC	346
Qy	868	ATGATGCGGGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	927
Db	347	ACGACGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	406
Qy	928	CGGATAACGCTTAATCTGGATAAAGCCCGCTGATTTGCTGGCCAGTCAAGAAAAATACG	987
Db	407	CCGACACGCGCAGCTTGGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG	466
Qy	988	GCTCCAGTATTTCTTGGGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1047
Db	467	GCAAGAGCTCTCATGGCGGAGCTGATTTTCGCGCGCAACTGCGCGCTGGAATCGA	526
Qy	1048	TGGGATTTAAAGCTGGGATTTTCGCGGAGAGAGAAAGTACTGGGAGTTCGGACCTGG	1107
Db	527	TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGCGCGGTCGACCATGGGAGCCGATGAG	586
Qy	1108	TATACTGGGGGCTTGACAAACAGCTCTTCGAGATAACCGGGATA---AAACGGGAAC	1164
Db	587	TCTATTGGGCAAGGAAGCCACTGGCTGCGGATGAGCGTTACAGCGGTAAGCGGGATC	646
Qy	1165	TTCAAGAACCTCTTGGCGCCACGACGATGGGACTTATTTATGTCATCTCGAAGGCCCGC	1224

647	TG	GGAGAACCCGCTGGCCGCGCGTGCAGATGGGCGTGATCTACGTGAACCCGGAGGGGCCGA	706
1225	QT	GTGGAAACACAGATCTCTGGCTTCGCGCGAAGATATCAGGAAGCTTTTTCACGTATGG	1284
707	DB	ACGGCAACCCGGACCCATCGCCGCGCGTGCACATTCGCGAGACGTTTCGGCGCATGG	766
1285	QT	CCATGGATGATGAGGACACTTGGCCCTGATCCGCGGAGGGCATACATTTGGTAAGCAC	1344
767	DB	CCATGAACGAGCTGGAACAGCGCGCTGATCGTGGCGGTACACTTTTCGGTAAGACCC	826
1345	QT	ATGTGTGACGCTCTCTGAATAATGATTGGCGCAGGGCGCTGATGGTGCACCTGTGGAGG	1404
827	DB	ATGGCGCG---GCCGCGCGATCTGGTCGCGCCCGAACCCGAGGCTGCTCGCTGGAGC	883
1405	QT	AGCAGGACTCGGATGGAAAAATAATGTGTACAGGAACCGCAAAATATACCATCACCA	1464
884	DB	AGATGGGGTTGGGCTGGAAGAGTGTGATGGCACCGAACCCGGTAAGGACCGGATCACCA	943
1465	QT	GTGGCCTGGAAGGACCTGGTCGA	1488
944	DB	CGCGCATCGAGGTCGTATGGACGA	967

RESULT 44

MTU41304	2225 bp	DNA	linear	BCT 07-JAN-1996
LOCUS	Mycobacterium tuberculosis H ₃₇ Rv	INH-resistant strain H0078/91, mutant		
DEFINITION	catalase-peroxidase (katG) gene, complete cds.			
ACCESSION	U41304			
VERSION	U41304.1	GI:1150701		

REFERENCE

REFERENCE	1 (bases 1 to 2225)
AUTHORS	Marttila, H. J., Soini, H., Huovinen, P. and Viljanen, M. K.
TITLE	katG gene mutations in isoniazid-resistant <i>Mycobacterium tuberculosis</i> strains isolated from Finnish patients
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2225)
AUTHORS	Marttila, H. J., Soini, H., Huovinen, P. and Viljanen, M. K.
TITLE	Direct Submission
JOURNAL	Submitted (27-NOV-1995) Harri J. Marttila, Mycobacterial Reference Laboratory, National Public Health Institute, Kiinamyllyankatu 13, Turku, SF-20520, Finland

FEATURES

source

gene

CDS

RESULT.T 45

RESULTS					PAT 07-OCT-1997
I61339	I61339	2235 bp	DNA	linear	
LOCUS					
DEFINITION	Sequence 1 from patent US 5658733.				
ACCESSION	I61339				
VERSION	I61339.1	GI:2479287			
KEYWORDS					

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2235)
TITLE	Cockerill, F.R. III, Kline, B.C. and Uhl, J.R.
JOURNAL	Detection of isoniazid resistant strains of M. tuberculosis
FEATURES	Patent: US 5658733-A 1 19-AUG-1997;
source	location/Qualifiers
	1..2235
BASE COUNT	431 a 696 c 745 g 363 t
ORIGIN	/organism="unknown"
Query Match	20.08; Score 298.4; DB 6; Length 2235;
Best Local Similarity	62.4%; Pred. No. 1.8e-77;
Matches 502; Conservative	0; Mismatches 296; Indels 6; Gaps 2;
Qy	688 ATCCCTGGGCGGTATTTTATTATGCCACAGATTTCAACAGCTCGATATGGAGCCTC 747
Db	
Qy	176 ACCGATGGGTGGCGCTTCGACTATGCCGGGAGTGCGACCATTGAGCTTGACGCC 235
Db	
Qy	748 TGA AAAAAGATATCAAAGATTGCTGCACAACTTCCCAGGATTTGGTGCCCTCGCGATTATG 807
Db	
Qy	236 TGACGCGGACATCGAGGAAGTAGTACCACCTCGCACGCCGTGTGTGSCCGGACTACG 295
Db	
Qy	808 GTCAATTATGTCCTTTCTTTATTCGTATGCGTTGGCAGCGTGCCGGAACAATACAGGACAT 867
Db	
Qy	296 GCCACTACGGGCGCGCTTTTATCCGATGGCGTGCCACGCTGCCGGCACCTACCGCATCC 355
Db	
Qy	868 ATGATGCCGGGAGGCCAGTGGTGGTCAGCAACGTTTTTGAACCGCTGAACAGCTGGC 927
Db	
Qy	356 ACGACGCCGCGCGCGCGCGCGCGATGACGCGTTGCGCCCGCTTAACAGCTGGC 415
Db	
Qy	928 CGGATAACGTTAATCTGGATAAAGCCCGTCGATTGCTGTGCCAGTCAAGAAAAAATACG 987
Db	
Qy	416 CCGAACACGCCAGCTTGTGAAGCGCGCGGCTGCTGTGCCGCTCAAGAAGAATGACG 475
Db	
Qy	988 GCTCCAGTATTCTCTGGGGAGACCTGATGCTCTGCTGCTGATGTTGCCCTTGAATCCA 1047
Db	
Qy	476 GCAAGAGCTCTATGCGCGGACCTGATGTTTTCGCGCGCACTCGCGCGTGGAAATCGA 535
Db	
Qy	1048 TGGGATTTAAACGCTGGGATTTGCTGCGGAGAGAAGATGACTGGGAGTCGACACCTGG 1107
Db	
Qy	536 TGGGCTTCAAGACGTTCCGGTTCCGGTTCCGCGGGTCGACCACTGGGAGCCCGATGAGG 595
Db	
Qy	1108 TATACTGGGGCCTGACAAACAGCCTCTTCGACATAACCGGATA----AAAACGGGAAC 1164
Db	
Qy	596 TCTATTGGGCAAGGAAGCACCACTGGCTCGCGATGACGCTTACAGCGGTAAAGCGGATC 655
Db	
Qy	1165 TTCAGAAACCTCTTGCCGCCACGAGATGGACATTTATTATGTCATCTCGAAGGCCCGG 1224
Db	
Qy	656 TGGNAAACCGCTGGCCCGGTGCAGATGGGGTGATCTACGTGAACCCCGAGGGGCCGA 715
Db	
Qy	1225 GTGAA AACAGATPCTCTTGCTTCCGCGAAGAGATATCAGGGAAGCTTTTTCAGTATGG 1284
Db	
Qy	716 ACGCCAACCCGACCCCATGGCCGCGCGTTCGACATTCGCGAGACGTTTCGGCGCATGG 775
Db	
Qy	1285 CCATGATGATAGGAACACTGTGGCCCTGATCGCGGAGGAGCATACATTTGGTPAAACAC 1344
Db	
Qy	776 CCATGACGAGCTCGA AACAGCGCGCTGATCGTCGCGGCTCACTTTTCGGTAAGACCC 835
Db	
Qy	1345 ATGGTCAGCGCTCTCTGAAAATGTATTGCGCAGGCGCTGATGTTGTCACCTCTGTGAGG 1404
Db	
Qy	836 ATGGCGCG---GCCCGCGCATCTGGTCGCCCGCAACCGAGGCTGTCTCCGCTGGAGC 892
Db	
Qy	1405 AGCAGGACCTGGGATGGA AAATAAATGTGTGTACAGGAACGGCAAAATATACCATCACCA 1464
Db	
Qy	893 AGATGGGCTGGGGCTGGAAGAGCTCGTATGSCACCGAACCGGTAAAGGACCGATACCA 952
Db	
Qy	1465 GTGCCCTGGAGGAGCCTGGTFCGA 1488
Db	
Qy	953 GCGGATCGAGGTGCTATGGACGA 976
Db	

LOCUS	2235 bp	DNA	linear	PAT 03-APR-1998
DEFINITION	Sequence 1 from patent US 5688639.			
ACCESSION	I74293			
VERSION	I74293.1	GI:3010434		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 2235)			
AUTHORS	Cockerill, F.R., Kline, B.C. and Uhl, J.R.			
TITLE	Detection of isoniazid resistant strains of <i>M. tuberculosis</i>			
JOURNAL	Patent: US 5688639-A 18-NOV-1997;			
FEATURES	Location/Qualifiers			
source	1..2235			
BASE COUNT	431 a	696 c	745 g	363 t
ORIGIN				
Query Match	20.0%	Score 298.4;	DB 6;	Length 2235;
Best Local Similarity	62.4%	Pred. No. 1.8e-77;		
Matches	502;	Conservative 0;	Mismatches 296;	Indels 6; Gaps 2;
Qy	688	ATCCCTGGGGGCTGATTTTGTATTATGCCACCAAGATTTCACACAGCTGGATATGGAGCTC	747	
Db	176	ACCGGATGGTGGCGCTTCGACTATGCCGGGAGGTGCGGACCATCGACGTTCAACGCC	235	
Qy	748	TGAAAAAGATATCAAGATTTCGTGACACTTCCCAGGATTTGGTGCCTTCGGGATTTATG	807	
Db	236	TGACGCGGGACATCGAGGAAAGTATGATACCACTCGCAGCGCTGTGTGCCCGCGGACTACG	295	
Qy	808	GTCATTATGTCTCTTCTTTATTCGTATGCTTGGCACGCTGCCGGAACATACAGGACAT	867	
Db	296	GCCACTACGGGCGCTGTTATCCGGATGGCTGGCACGCTGCCGGCACCTACCGCATCC	335	
Qy	868	ATGATGCCGGGAGGCCAGTGGTGTCAGCAACGTTTTGAACCCCTGAACAGCTGGC	927	
Db	356	ACGACGCCGCGGCGCGCGGCGGCGCATGACGCGTTTCGGCGCCTTTAACAGCTGGC	415	
Qy	928	CGGATACGTTAATCTGGATTAAGCCCGTCTGATTGCTGTGCCAGTCAAGAAAAATACG	987	
Db	416	CCGACAACGCCGCTTGGCAAGCGCGCGGCTGTGTGCGCGGTCAAGAAGAAGTACG	475	
Qy	988	GCTCCAGTATTTCTTGGGAGACCTGATGTCCTGACTGCTAATGTTGCCCTTGAATCCA	1047	
Db	476	GCAAGAGCTCTCATGGCGGACCTGATGTTTTCGCGGCACTGCGCCTGGATCTGA	535	
Qy	1048	TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAAGATGATGSGAGCTCGGACCTGG	1107	
Db	536	TGGGCTTCAAGACGTTCCGGTTCGGCTTCGGCGGGTCGACCACTGGGAGCCCGATGAGG	595	
Qy	1108	TATACTGGGGCTTGACAAACAGCCTCTTGCAGATAACCGGGATA---AAACGGGAAC	1164	
Db	596	TCTATTGGGCAAGGAACCAACCTTGGCTCGCGCATGAGCGTTTACAGCGGTAAAGCGGATC	655	
Qy	1165	TTCAGAAACCTTTTGGCGCCACGACGATGGGACTTATTTATGTCAATCTTGAAGGCCCGC	1224	
Db	656	TGGAGAACCGCTGCGGCGCGGTGCAGATGGGCTGTATCTACGTAAACCCGAGGGGCGGA	715	
Qy	1225	GTGAAACACAGATCTCTTGCTTCCCGGAAAGATATCAGGGAAGCTTTTTCACGTATGG	1284	
Db	716	ACGCAACCCGGACCCCATGGCGCGCGCTGCACATTCGGAGACGTTTCGGGGCATGG	775	
Qy	1285	CCATGTGATGATGAGGAGACTGTGGCCCTGTATCGCGGAGGGGCATACATTTGGTAAAGCAC	1344	
Db	776	CCATGAACGAGTCTGAAACACGCGCGCTGATCGTGGCGGTTCACACTTTCGGTAAAGACC	835	
Qy	1345	ATGGTCAGCGTCTCTTGAAATAATGTATTTGGCGGACGGCCCTGTATGTGACCTGTGGAGG	1404	
Db	836	ATGGCGCG---GCCCGGCGGATCTGTGTGCGCCCGCAACCCGAGGCTGCTCCGCTGGAGC	892	

QY	1108	TATACTGGGGGCTGACAAACAGCCCTCTTGACAGATAACCGGGATA---	AAAAACGGAAAC	1116
Db	636	TCATTATGGGCAAGGAAGCCACCTGGCTCGGCATGAGCGTTACAGCGGTAAACGGGATC	695	
QY	1165	TTCAGAAACCTCTTGGCCGACACAGATGGGACTTATTTATGTCAATCTCTGAAGGCCCG	1224	
Db	696	TGAGAACCCGCTGGCCGGGTGCAGATGGGCTGATCTACGTGAACCCGAGGGGCCGA	755	
QY	1225	GTGGAACACCATCTCTGGCTTCGCGAAGATATCAGGGAAGCTTTTTCACGTATGG	1284	
Db	756	ACGGCAACCCGGACCCATGGCCGGCGGTGCACATTCGCGAGAGCTTTCGCGCATGG	815	
QY	1285	CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCAATATTTGGTAAAGCAC	1344	
Db	816	CCATGAACGACGTCGAAACAGCGCGCTGATCTCGCGGCTCACACTTTCGGTAAGACCC	875	
QY	1345	ATGGTCGACGCTCTCTGAAATATGATTGGCCAGGGCCTGATGTGACCTGTGGAGG	1404	
Db	876	ATGGCCCG---GCCCGCGCATCTGGTCGGCCCGAACCCGAGGCTGCTCCCTCGAGC	932	
QY	1405	AGCAGGACTGGGATGGAATAAATATGTGTACAGGAACGCAATATACCATCACCA	1464	
Db	933	AGATGGCTGGGCTGGAGAGCTGATGGCACCGGACCGGTAAGGACCGCATCACCA	992	
QY	1465	GTGGCTGGAAGGAGCTGTGCGA	1488	
Db	993	GCGGATCGAGGTCGTATGGACGA	1016	
RESULT 49				
MTU06264				
LOCUS				
DEFINITION		2316 bp DNA linear BCT 20-MAY-1994		
ACCESSION		Mycobacterium tuberculosis L11150 catalase (katG) gene, complete		
VERSION		D06264		
KEYWORDS		U06264.1 GI:488443		
SOURCE				
ORGANISM		Mycobacterium tuberculosis.		
REFERENCE		Mycobacterium tuberculosis.		
AUTHORS		Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
TITLE		Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium		
JOURNAL		tuberculosis complex.		
REFERENCE		1 (sites)		
AUTHORS		Cockerill, F.R., Uhl, J.R., Temesgen, Z., Zhang, Y., Stockman, L.,		
TITLE		Roberts, G.D., Williams, D.L. and Kline, B.C.		
JOURNAL		Rapid identification of a point mutation of the Mycobacterium		
AUTHORS		tuberculosis catalase-peroxidase (katG) gene associated with		
TITLE		isoniazid resistance		
JOURNAL		Unpublished		
REFERENCE		2 (bases 1 to 2316)		
AUTHORS		Kline, B.C.		
TITLE		Direct Submission		
JOURNAL		Submitted (07-FEB-1994) Bruce C. Kline, Mayo Clinic and Mayo		
FEATURES		Foundation, Biochemistry and Molecular Biology, 200 SW 1 St.,		
source		Rochester, MN 55905, USA		
gene		Location/Qualifiers		
CDS		1..2316		
		/organism="Mycobacterium tuberculosis"		
		/strain="L11150"		
		/db_xref="taxon:1773"		
		69..2291		
		/gene="katG"		
		69..2291		
		/gene="katG"		
		/standard_name="catalase-peroxidase"		
		/codon_start=1		
		/transl_table=11		
		/product="catalase"		
		/protein_id="AA18234.1"		
		/db_xref="GI:488443"		
		/translation="MPEQHPPIETTTGAASNCSPVVGHMKYVPEGGGNDDWWPNRLN		
		LKYLHQAQVADPMDGAADYAAAVATIDVALTRDIEVNTSQPWPADYGHGFLP		
		IRNMAHQACTYRIHDGAGGAGGQRFAPLNSWPNDSLDKARLLWPKVKYKKGLS		

WADLVFAGNCALESMTGFTGFGFGRVDQWEPDEYWGKEATWLGDERYSKRDLEN
PLAAYQMLIVYVPEGNPNPDMAAAVDIRETFRAMNDVETAAALIVGGHTFGKTH
GAGPADLVGPEPEAALEQMLGWSYGTGKDAITSGIEVNTNTTWDNSELE
ILYVYEWELTKSPAGAWOYAKDAGAGATIPDPFGGPRSPMLATDLSLRVDPYIER
ITRWLEHPEELADEFAKAWYKLIHRDMGPVARYLGLVPKOTLLWQDVPVASHDLV
GEAEIASLKSQIRASGLTYSQVSTAWAASRSGRSGKGGANGRRILQPOGWEN
DPDGLRKVIRTLLEIESNSAAGNIKRVFADLVVLGGCAIEKAAKAAGHNITVP
FTFGRDASQEDVESFAVLEPKADGFRNGLGNPLPAEYMLDKANLLLSAPEM
TVLVGLRLVGLANYKRLPLGVFTFESLSLNDFFVNLDMGITWEPSPADGGTYOGKD
GSGKVKWTGSRVLDVFGSSELRAVVEVYTGADDAQPKFQVDFVAWMDKVNLDREDFVR

BASE COUNT	449 a	725 c	760 g	378 t	3 others
ORIGIN					
Query Match	20.0%; Score 298.4; DB 1; Length 2316;				
Best Local Similarity	62.4%; Pred. No. 1.8e-77;				
Matches	502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;				
QY	688	ATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATATGAGGCTC	747		
DB	235	ACCCGATGGGTGGCGGCTTGACCTATCGCGGAGGTGCGACCATCGAGCTTGAGCGCC	294		
QY	748	TGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGTCCTGCGGATATG	807		
DB	295	TGACCGGGACATCGAGAGTATGACCACTCGACCGCTGCTGGCCCGCGACTACG	354		
QY	808	GTCATTATGTCCTTTCTTTATTCGTATGCTTTGGCACGCTGGCGGAACATACAGACAT	867		
DB	355	GCCACTACGGCGGCTCTTTATCGGATGCGTGGCACGCTGCGGCACTACCGCATCC	414		
QY	868	ATGATGCGCGGGAGCGGAGTGGTGGTGCAGCAAGCTTTTGAACCGCTGAACAGTGGC	927		
DB	415	ACGACGCGCGGCGCGCGCGGCGGCGATGCGAGCGGTTTTCGCGCGCTTAAACAGTGGC	474		
QY	928	CGGATAACGTTAATCGATAAGCCGCTGATGCTGTCGCGGATGCTGCGCAAGAAAAATACG	987		
DB	475	CCGACACGCCAGCTTGACAGCGCGCGGCTGCTGTCGCGGCTCAAGAAAGTACG	534		
QY	988	GCTCCAGTATTCCTGGGAGACCTGATGTCCTGACTGTGTAATGTTGCCCTTGAATCCA	1047		
DB	535	GCAAGAAGCTCTACGCGGCGGACCTGATGTTTTCGCGGCACTGCGGCTGGAATCGA	594		
QY	1048	TGGGATTTAAACCTGGGATTTGCTGGCGGAGAGAGATGATCGGAGTGGACCTGG	1107		
DB	595	TGGGCTTCAAGACCTTCGGGTTGCGCTTCGCGGCGGCTGACACGATGGAGCCCGATGAGG	654		
QY	1108	TATACTGGGCGCTTGACAAAGCCTCTTCAGATTAACCGGGATA---AAAAAGGAAAC	1164		
DB	655	TCTATTGGGCAAGAACCCACCTGCTGCGGATGAGCTTACAGCGGTAAAGCGGATC	714		
QY	1165	TTCAGAACCTCTTTGCCCGCACGAGATGGAGCTTATTTATGCAATCTTGAAGGCCCG	1224		
DB	715	TGGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGAGGGGCGCA	774		
QY	1225	GTGAAACACGATCTCTGCTTCCCGGAAAGATATCAGGAAGCTTTTTCACGTATGG	1284		
DB	775	ACGGCAACCCGAGCCCATGCGCGCGGCTGCGACATTCGCGGAGAGCTTTTCGCGCGATGG	834		
QY	1285	CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC	1344		
DB	835	CCATGACGAGCTGCAACAGCGCGCTGATCTGCGCGGTACACATTTTCGGTAAAGCCC	894		
QY	1345	ATGGTGACGCTCTCTTCAAAAATGATTTGGCCGAGGCGCTGATGTCACCTGTGGAGG	1404		
DB	895	ATGGCGCG---GCCCGCGCATCTGCTGGCGCCCGAACCCGAGGCTGCTCCGCTGGAGC	951		
QY	1405	ACGAGGACTGGGATGAAATAAATGTTGGTACAGGAACGGCAATATACCATCCCA	1464		
DB	952	AGATGGGCTGGGTGGAAGAGCTCGTATGGCACCGAACCGGTGAAGGACGCGATCACC	1011		
QY	1465	GTGGCTGGAAGGAGCTTGGTCA	1488		
DB	1012	CGGCACTGAGGCTGTTGGACGA	1035		

RESULT 50
MTU06260
LOCUS
DEFINITION

MTU06260 2326 bp DNA linear BCT 20-MAY-1994
Mycobacterium tuberculosis G6108 catalase (katG) gene, complete cds.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

U06260
U06260.1 GI:488435

Mycobacterium tuberculosis.

Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

1..2326

/organism="Mycobacterium tuberculosis"

/strain="G6108"

/db_xref="taxon:1773"

70..2292

/gene="katG"

70..2292

/gene="katG"

/function="catalase-peroxidase"

/codon_start=1

/transl_table=11

/product="catalase"

/protein_id="AA18231.1"

/db_xref="GI:488435"

/translation="MPEQHPITETTTGAASNGCPVGHMKYVPGGGNQDWWNRLN

LKVLQHNQAVADPMGAADYAAEVATIDVADLTREIVMTTSQPMWADYGHYGLF

IRMAHAGTYRIHDGCGGQMRFAPLSNPDNAISDKARLLWPKYKYGKLS

WDLIVFAGNCALESMTGFTGFGFGRVDQWEPDEYWGKEATWLGDERYSKRDLEN

PLAAYQMLIVYVPEGNPNPDMAAAVDIRETFRAMNDVETAAALIVGGHTFGKTH

ITRWLEHPEELADEFAKAWYKLIHRDMGPVARYLGLVPKOTLLWQDVPVASHDLV

GEAEIASLKSQIRASGLTYSQVSTAWAASRSGRSGKGGANGRRILQPOGWEN

DPDGLRKVIRTLLEIESNSAAGNIKRVFADLVVLGGCAIEKAAKAAGHNITVP

FTFGRDASQEDVESFAVLEPKADGFRNGLGNPLPAEYMLDKANLLLSAPEM

TVLVGLRLVGLANYKRLPLGVFTFESLSLNDFFVNLDMGITWEPSPADGGTYOGKD

GSGKVKWTGSRVLDVFGSSELRAVVEVYTGADDAQPKFQVDFVAWMDKVNLDREDFVR

"

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

502; Conservative

0; Mismatches

296; Indels

6; Gaps

2;

QY

688

ATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATATGAGGCTC

747

DB

236

ACCGATGGGTGGCGGCTTGACCTATCGCGGAGGTGCGGACCATCGAGCTTGAGCGCC

295

QY

748

TGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGCGGATATG

807

DB

296

TGACCGGGACATCGAGGAAGTATGATCACCACCTCGCAGCGCTGGTGGCCCGCGACTACG

355

QY

808

GTCATTATGGTCTCTTTTTCGTTATGCTATGGCTGGCACCGTGGCCGGAACATACAGACAT

867

Db 356 GCCACTACGGCGCGCTTTATCCGGATGCGTGGACACGCTCGCGGACCTACCGATCC 415
Qy 868 ATGATGCCCGGGAGGCGCCAGTGGTGCAGCAACGTTTGAACCGCTGAACAGCTGGC 927
Db 416 ACACGCGCGCGCGCGCGCGCGCATGCGCGGTTCCGCGCGCTTAACAGCTGGC 475
Qy 928 CGGATACGTTAATCTGGATAAAGCCCGTCGATGCTGTGGCCAGTCAAGAAAATACG 987
Db 476 CCGACAAACGAGCTTGGACAAGCGCGCGCGCTGTGTGGCCGGTCAAGAAGATACG 535
Qy 988 GCTCCAGTATTCTGGGAGACCTGATGCTCTGACTGTAATGTTGCCCTTGAATCCA 1047
Db 536 GCAGAGAGCTCTATGGCGGAGCTGATTTTCGCCGCACTGCGCGCTGGATCGA 595
Qy 1048 TGGGATTTAAACCGCTGGGATTTCTGGCGGAGAGAAAGATGACTGGGAGTTCGGACCTGG 1107
Db 596 TGGGCTTCAAGAGCTTCGGGTTTCGGCTTCGGCGGGTCGACCACTAGTGGGAGCCGATGAGG 655
Qy 1108 TATACTGGGGCGCTGACAAAGCCCTTTCAGATACCCGGGATA ---AAAACGGGAAC 1164
Db 656 TCTATTGGGCAAGAGAGACCCCTGGCTCGCGCATGAGCGTTACAGCGGTAAGCGGATC 715
Qy 1165 TTCAGAAACCTCTTGC CGCCACGACATGGACTTATTTATGTCAATCCTGAAGGCCCG 1224
Db 716 TGGAGAACCGCTGGCGCGCGGTGCAGATGGGGTGTACTACGTGAACCCGAGGGCGCGA 775
Qy 1225 GTGGAAACAGATCCTCTCGCTTCCGCGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
Db 776 ACGCAACCGGACCCCATGGCCGCGCGGTGCAGCATTCGCGAGACGTTTCGGCGCATGG 835
Qy 1285 CCATGGATGATGAGGAGACTGTGGCCCTGATCGGGGAGGGCATACATTTGGTAAAGCAC 1344
Db 836 CCATGAACGACGTGCGAAACAGCGCGCGCTGATCGGCGGTACACACTTTCGGTAAAGACC 895
Qy 1345 ATGGTCAGCGCTCTCCTGAAAATGTATTGGCGCAGGGCCCTGATGGTCACCTGTGGAGG 1404
Db 896 ATGGCGCG ---GCCCGCGCATCTGGTCGCGCCCGAACCAGGCTGCTCCGCTGGAGC 952
Qy 1405 AGCAGGAGCTGGGATGGAAAAATAAATGTGGTACAGGAAACGGCAAAATATACCATCACCA 1464
Db 953 AGATGGGCTTGGGCTGGAGAGCTCGTATGGCACCGGACCGGTAAAGGACGCGATCACCA 1012
Qy 1465 GTGGCCTGGAAGGAGCCTGGTCGA 1488
Db 1013 GCGGCATCGAGGTCGTATGGACGA 1036

Search completed: June 18, 2003, 04:52:44

Job time : 2697.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 04:53:09 ; Search time 376.433 Seconds
(without alignments)
5804.490 Million cell updates/sec

Title: US-09-674-277-1

Perfect score: 1489

Sequence: 1 ctgcagtcgagatgaaag.....ctggaagagcctgctgac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357.6	24.0	2238	10	US-09-884-889-7
2	224.6	15.1	2262	10	US-09-884-889-5
3	38.6	2.6	567	9	US-10-123-155-254
4	34.6	2.3	520	9	US-10-184-634-332
5	34.6	2.3	520	9	US-10-184-634-332
6	34.2	2.3	893	10	US-09-815-242-2188
7	34.2	2.3	927	10	US-09-815-242-8989
8	34.2	2.3	1164	10	US-09-815-242-4832
9	34.2	2.3	2354	7	US-08-781-986A-565
10	34	2.3	636	9	US-09-764-891-8918
11	33.8	2.3	457	10	US-09-864-761-3110
12	33.8	2.3	499	9	US-10-184-644-592
13	33.8	2.3	499	9	US-10-184-634-592
14	33.8	2.3	579	9	US-09-924-400-147
15	33.8	2.3	579	10	US-09-810-936-147
16	33.8	2.3	579	10	US-09-429-755-147
17	33.8	2.3	882	9	US-10-184-644-574
18	33.8	2.3	882	9	US-10-184-634-574
19	33.4	2.2	6383	9	US-10-092-154-1415

c	20	33.4	2.2	6383	10	US-09-764-847-1415	Sequence 1415, Ap
c	21	33.4	2.2	21732	9	US-10-092-154-1414	Sequence 1414, Ap
c	22	33.4	2.2	21732	10	US-09-764-847-1414	Sequence 1414, Ap
	23	33.2	2.2	484	9	US-09-918-995-29551	Sequence 29551, A
	24	33.2	2.2	3226	9	US-09-992-598-352	Sequence 352, App
	25	33.2	2.2	3226	9	US-09-989-293A-352	Sequence 352, App
	26	33.2	2.2	3226	9	US-09-989-733-352	Sequence 352, App
	27	33.2	2.2	3226	9	US-09-990-444-352	Sequence 352, App
	28	33.2	2.2	3226	9	US-09-989-730-352	Sequence 352, App
	29	33.2	2.2	3226	9	US-09-990-436-352	Sequence 352, App
	30	33.2	2.2	3226	9	US-09-991-181-352	Sequence 352, App
	31	33.2	2.2	3226	9	US-09-993-687-352	Sequence 352, App
	32	33.2	2.2	3226	9	US-09-989-734-352	Sequence 352, App
	33	33.2	2.2	3226	9	US-10-028-072-463	Sequence 463, App
	34	33.2	2.2	3226	9	US-09-997-653-352	Sequence 352, App
	35	33.2	2.2	3226	9	US-09-993-667-352	Sequence 352, App
	36	33.2	2.2	3226	9	US-10-121-049-463	Sequence 463, App
	37	33.2	2.2	3226	9	US-10-123-904-463	Sequence 463, App
	38	33.2	2.2	3226	9	US-10-140-470-463	Sequence 463, App
	39	33.2	2.2	3226	9	US-09-990-438-352	Sequence 352, App
	40	33.2	2.2	3226	9	US-09-990-562-352	Sequence 352, App
	41	33.2	2.2	3226	9	US-09-997-428-352	Sequence 352, App
	42	33.2	2.2	3226	9	US-09-997-666-352	Sequence 352, App
	43	33.2	2.2	3226	9	US-10-175-746-463	Sequence 463, App
	44	33.2	2.2	3226	9	US-10-176-918-463	Sequence 463, App
	45	33.2	2.2	3226	9	US-10-176-921-463	Sequence 463, App
	46	33.2	2.2	3226	9	US-09-990-711-352	Sequence 352, App
	47	33.2	2.2	3226	9	US-10-137-865-463	Sequence 463, App
	48	33.2	2.2	3226	9	US-10-140-474-463	Sequence 463, App
	49	33.2	2.2	3226	9	US-10-142-431-463	Sequence 463, App
	50	33.2	2.2	3226	9	US-10-143-114-463	Sequence 463, App
	51	33.2	2.2	3226	9	US-09-989-726-352	Sequence 352, App
	52	33.2	2.2	3226	9	US-10-140-002-463	Sequence 463, App
	53	33.2	2.2	3226	9	US-09-990-437-352	Sequence 352, App
	54	33.2	2.2	3226	9	US-09-998-156-352	Sequence 352, App
	55	33.2	2.2	3226	9	US-10-142-419-463	Sequence 463, App
	56	33.2	2.2	3226	9	US-09-991-157-352	Sequence 352, App
	57	33.2	2.2	3226	9	US-09-991-172-352	Sequence 352, App
	58	33.2	2.2	3226	9	US-09-997-514-352	Sequence 352, App
	59	33.2	2.2	3226	9	US-09-997-573-352	Sequence 352, App
	60	33.2	2.2	3226	9	US-10-123-262-463	Sequence 463, App
	61	33.2	2.2	3226	9	US-10-142-423-463	Sequence 463, App
	62	33.2	2.2	3226	9	US-09-990-443-352	Sequence 352, App
	63	33.2	2.2	3226	9	US-09-990-726-352	Sequence 352, App
	64	33.2	2.2	3226	9	US-09-997-559-352	Sequence 352, App
	65	33.2	2.2	3226	9	US-09-997-601-352	Sequence 352, App
	66	33.2	2.2	3226	9	US-10-121-050-463	Sequence 463, App
	67	33.2	2.2	3226	9	US-10-141-755-463	Sequence 463, App
	68	33.2	2.2	3226	9	US-09-989-729A-352	Sequence 352, App
	69	33.2	2.2	3226	9	US-09-990-440-352	Sequence 352, App
	70	33.2	2.2	3226	9	US-09-991-854-352	Sequence 352, App
	71	33.2	2.2	3226	9	US-09-997-349-352	Sequence 352, App
	72	33.2	2.2	3226	9	US-09-997-440-352	Sequence 352, App
	73	33.2	2.2	3226	9	US-09-997-628-352	Sequence 352, App
	74	33.2	2.2	3226	9	US-09-997-683-352	Sequence 352, App
	75	33.2	2.2	3226	9	US-10-143-032-463	Sequence 463, App
	76	33.2	2.2	3226	9	US-09-993-469-352	Sequence 352, App
	77	33.2	2.2	3226	9	US-09-993-748-352	Sequence 352, App
	78	33.2	2.2	3226	9	US-09-997-542-352	Sequence 352, App
	79	33.2	2.2	3226	9	US-10-123-108-463	Sequence 463, App
	80	33.2	2.2	3226	9	US-10-123-236-463	Sequence 463, App
	81	33.2	2.2	3226	9	US-10-123-261-463	Sequence 463, App
	82	33.2	2.2	3226	9	US-10-140-921-463	Sequence 463, App
	83	33.2	2.2	3226	9	US-10-140-928-463	Sequence 463, App
	84	33.2	2.2	3226	9	US-09-990-427-352	Sequence 352, App
	85	33.2	2.2	3226	9	US-09-990-439-352	Sequence 352, App
	86	33.2	2.2	3226	9	US-10-121-045-463	Sequence 463, App
	87	33.2	2.2	3226	9	US-10-123-292-463	Sequence 463, App
	88	33.2	2.2	3226	9	US-10-123-903-463	Sequence 463, App
	89	33.2	2.2	3226	9	US-10-124-819-463	Sequence 463, App
	90	33.2	2.2	3226	9	US-10-124-822-463	Sequence 463, App
	91	33.2	2.2	3226	9	US-10-140-925-463	Sequence 463, App
	92	33.2	2.2	3226	9	US-10-160-498-463	Sequence 463, App

93	33.2	2.2	3226	9	US-09-989-328-352	Sequence 352, App
94	33.2	2.2	3226	9	US-09-993-583-352	Sequence 352, App
95	33.2	2.2	3226	9	US-10-121-041-463	Sequence 463, App
96	33.2	2.2	3226	9	US-10-121-043-463	Sequence 463, App
97	33.2	2.2	3226	9	US-10-121-047-463	Sequence 463, App
98	33.2	2.2	3226	9	US-10-123-215-463	Sequence 463, App
99	33.2	2.2	3226	9	US-10-123-902-463	Sequence 463, App
100	33.2	2.2	3226	9	US-10-123-908-463	Sequence 463, App
101	33.2	2.2	3226	9	US-10-123-909-463	Sequence 463, App
102	33.2	2.2	3226	9	US-10-123-910-463	Sequence 463, App
103	33.2	2.2	3226	9	US-10-124-813-463	Sequence 463, App
104	33.2	2.2	3226	9	US-10-124-817-463	Sequence 463, App
105	33.2	2.2	3226	9	US-10-124-824-463	Sequence 463, App
106	33.2	2.2	3226	9	US-10-125-922-463	Sequence 463, App
107	33.2	2.2	3226	9	US-10-125-924-463	Sequence 463, App
108	33.2	2.2	3226	9	US-10-127-825A-463	Sequence 463, App
109	33.2	2.2	3226	9	US-10-127-829A-463	Sequence 463, App
110	33.2	2.2	3226	9	US-10-127-835A-463	Sequence 463, App
111	33.2	2.2	3226	9	US-10-127-839A-463	Sequence 463, App
112	33.2	2.2	3226	9	US-10-127-901A-463	Sequence 463, App
113	33.2	2.2	3226	9	US-10-128-693A-463	Sequence 463, App
114	33.2	2.2	3226	9	US-10-131-813A-463	Sequence 463, App
115	33.2	2.2	3226	9	US-10-131-818A-463	Sequence 463, App
116	33.2	2.2	3226	9	US-10-131-823A-463	Sequence 463, App
117	33.2	2.2	3226	9	US-10-131-824A-463	Sequence 463, App
118	33.2	2.2	3226	9	US-10-131-830A-463	Sequence 463, App
119	33.2	2.2	3226	9	US-10-131-837A-463	Sequence 463, App
120	33.2	2.2	3226	9	US-10-137-872A-463	Sequence 463, App
121	33.2	2.2	3226	9	US-10-140-860-463	Sequence 463, App
122	33.2	2.2	3226	9	US-10-142-417-463	Sequence 463, App
123	33.2	2.2	3226	9	US-10-147-500-463	Sequence 463, App
124	33.2	2.2	3226	9	US-10-147-502-463	Sequence 463, App
125	33.2	2.2	3226	9	US-10-147-515-463	Sequence 463, App
126	33.2	2.2	3226	9	US-10-147-517-463	Sequence 463, App
127	33.2	2.2	3226	9	US-10-147-519-463	Sequence 463, App
128	33.2	2.2	3226	9	US-10-147-526-463	Sequence 463, App
129	33.2	2.2	3226	9	US-10-147-537-463	Sequence 463, App
130	33.2	2.2	3226	9	US-10-153-395-463	Sequence 463, App
131	33.2	2.2	3226	9	US-10-157-792-463	Sequence 463, App
132	33.2	2.2	3226	9	US-09-941-992-352	Sequence 352, App
133	33.2	2.2	3226	9	US-09-992-521-352	Sequence 352, App
134	33.2	2.2	3226	9	US-10-121-040-463	Sequence 463, App
135	33.2	2.2	3226	9	US-10-121-056-463	Sequence 463, App
136	33.2	2.2	3226	9	US-10-121-061-463	Sequence 463, App
137	33.2	2.2	3226	9	US-10-123-235-463	Sequence 463, App
138	33.2	2.2	3226	9	US-10-124-818-463	Sequence 463, App
139	33.2	2.2	3226	9	US-10-125-926A-463	Sequence 463, App
140	33.2	2.2	3226	9	US-10-125-930A-463	Sequence 463, App
141	33.2	2.2	3226	9	US-10-127-831A-463	Sequence 463, App
142	33.2	2.2	3226	9	US-10-127-837A-463	Sequence 463, App
143	33.2	2.2	3226	9	US-10-127-838B-463	Sequence 463, App
144	33.2	2.2	3226	9	US-10-127-842A-463	Sequence 463, App
145	33.2	2.2	3226	9	US-10-127-843A-463	Sequence 463, App

ALIGNMENTS

RESULT 1

US-09-884-889-7
 ; Sequence 7, Application US/09884889
 ; Patent No. US20020102680A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: ROBERTSON, Dan
 ; APPLICANT: SANVAL, Indrajit
 ; APPLICANT: ADHIKARI, Robert
 ; TITLE OF INVENTION: CATALASES


```

/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2188
/ LENGTH: 893
/ TYPE: DNA
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-2188

```

	Query Match Best Local Matches	Similarity 49.7%; Conservative 87;	Score 34.2; Pred. No. 2; Mismatches 0;	DB 10; Indels 88;	Length 893; Gaps 0;
QY	483	AAAGTTGACATTTCTATACAAATGCGAAATTTCCCTTAATCCGGAGCTATTCGTATGATA	542		
Db	30	AAAGTTCAAATGCTTCTGCAATCTTGGAAATTTGCCCTTTTTATCAGCGAACTCATCTAATA	89		
QY	543	AAAAAAACTCTTCTCTGTCTGATTTCTTCTGGCGCTATCGGGAGCTTTTTCTACCGCGTGTA	602		
Db	90	ATAATTTTTTTTAAATTTATCTGGTCTCTTCAGTAGGACTTTGCACGATTTGAAAAACTATACG	149		
QY	603	CGCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAA	657		
Db	150	TCGGCTCTAGTACAAATTTGATACCAAAATGTATCATCGATAAGTCTTCGTGTTAA	204		

```

RESULT 7
/
/ SEQUENCE 8989, Application US/09815242
/ Patent No. US20020061569A1
/
/ GENERAL INFORMATION:
/
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/
/ TITLE OF INVENTION: Identification of
/ TITLE OF INVENTION: Prokaryotes
/
/ FILE REFERENCE: ELTRA 011A
/
/ CURRENT APPLICATION NUMBER: US/09/815
/ CURRENT FILING DATE: 2001-03-21
/
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/
/ NUMBER OF SEQ ID NOS: 14110
/
/ SOFTWARE: FASTSEQ for Windows Version
/ SEQ ID NO 8989

```

```

; LENGTH: 927
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(927)
US-09-815-242-8989

Query Match      2.3%:   Score 34.2
Best Local Similarity 49.7%; Pred. No.
Matches 87; Conservative 0; Mismatch

Qy 483 AAAGTTGACTTTTGTATACACATGCGGAATG
      ||||| | | | | | | | | | | | |
Db 715 AAAGTTCAAATGCTTCTGCATCTTGGAAATG
      ||||| | | | | | | | | | | | |
Qy 543 AAAAANAATCTTCTGTTTCTGATCTTCTG
      || | | | | | | | | | | | | |
Db 655 ATAATTTTTTAAATTTATCTGGTCTTCTCAG
      || | | | | | | | | | | | | |
Qy 603 GCCGCTGATAAAAAAGACATCAAAATTTTC
      || | | | | | | | | | | | | |
Db 595 TCGGCTCTAGTACAAATTTGATAACCAATG

RESULT 8
US-09-815-242-4832/c
; Sequence 4832, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyaskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Ess
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4832
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4832

```

	Query Match	2.38;	Score 34.2;	DB 10;	Length 1164;
	Best Local Similarity	49.7%;	Pred. No. 2.5;		
	Matches 87;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;
QY	483	AAAGTTGACCTTTTGTATACAACTGCGAAATTTCCCTTAATCCGGAGCTATTTCGTATGATA	542		
DB	1021	AAAGTTCAAATGCTTTCGTATCTTGGAAATTCCTTTTTATCAGCAACTCATCTAATA	962		

NOT RECENT TRAVEL:

PRIOR FILING DATE: 2000-09-21

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3110
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035457.11
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
; US-09-864-761-3110

Query Match      2.3%; Score 33.8; DB 10; Length 457;
Best Local Similarity 56.9%; Pred. No. 1.7;
Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1290 GATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGGATACATTTGGTAAGCACATGGT 1349
Db 219 GCTGCTGATGTTAAAGCAGGTGTGATGCTGATGAGGTACCGATGCTGAAGCAGGTGTT 160

QY 1350 GCAGCGTCTCTCTGAAAATATATTGGCGGAGGCGCTGATGTCACCTG 1398
Db 159 GATGACGCTCTGAAAAGGTGTGTTGGTGCAGGTGCAGATGCTGATGCTG 111

RESULT 12
US-10-184-644-592
; Sequence 592, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 592
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-592

Query Match      2.3%; Score 33.8; DB 9; Length 499;
Best Local Similarity 6.6%; Pred. No. 1.8;
Matches 32; Conservative 156; Mismatches 298; Indels 0; Gaps 0;

QY 681 GAATCAATCCCTGGGGGCTGATTTGATATGCCACACAGATTTCAACAGCTGGATATG 740
Db 9 WLLVLRVPRVPGQLDPSTGRRFSEHKLCADDECSMMYRGEALDFTGDCRFVNFKG 68

QY 741 GAGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGCG 800
Db 69 DPVYVYKLRGPEWAGSVGRFTGFPKDLIQVHVEYKKEELQVPTDETDFVCFDGR 128

QY 801 GATTATGCTATATGCTCTCTCTTTATTCGATGCTTGGCAGCTGCGCGGACATAC 860

```

```

QY 801 GATTATGCTATATGCTCTCTCTTTATTCGATGCTTGGCAGCTGCGCGGACATAC 860
Db 129 DDFHYNVEELGLFLELYNSAATSEKAVEKTLQDMKKNPELSKERPEPEPEANSEES 188
QY 861 AGGACATATGATGCGCGGAGCGCGAGTGGTGTCAGCAACGTTTGAACCGCTGAAC 920
Db 189 DSVFSENTEDLQOFTQTKRHHSHANSQANHAQEQASFSFEMLQDKLVPESENKTS 248
QY 921 AGCTGCCGGATAACGTTAATCTGGATAAAGCCGCTGATTTGCTGCTGCCAGTCAAGAA 980
Db 249 NSQVSNQDKIDAYKLLKEMTLDLTKTFGSTADALVSDDETRLTSLDEDDFDELDT 308
QY 981 AAATACGGTCCAGTATTTCTCTGGGAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1040
Db 309 EYAVGKEDENQEDFDELPLLTFTDGEDMKTFKSGVERKYPTDKQNSNEEDKVLQVTP 368
QY 1041 GAATCCATGGGATTTAAACGCTGGGATTTCTGGCGGAGAGAGAGATGATGCGGATCG 1100
Db 369 PGINKNDKNILTTWGTDTIFSVITGGETRTMDLESSSEEEKEDDDALVPSKQKQPQ 428
QY 1101 GACCTGTATPACTGGGGGCTGACAAACAGCTCTTTCAGATAACCGGATAAAACGGG 1160
Db 429 SATDYSDDPNVDGLFVDPKTNNDKEVNAEHKIKGRGVQESKRGVLQVDETELEDEN 488
QY 1161 AAACCTT 1166
Db 489 QEGFKT 494

RESULT 13
US-10-184-634-592
; Sequence 592, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 592
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-592

Query Match      2.3%; Score 33.8; DB 9; Length 499;
Best Local Similarity 6.6%; Pred. No. 1.8;
Matches 32; Conservative 156; Mismatches 298; Indels 0; Gaps 0;

QY 681 GAATCAATCCCTGGGGGCTGATTTGATATGCCACACAGATTTCAACAGCTGGATATG 740
Db 9 WLLVLRVPRVPGQLDPSTGRRFSEHKLCADDECSMMYRGEALDFTGDCRFVNFKG 68

QY 741 GAGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGCG 800
Db 69 DPVYVYKLRGPEWAGSVGRFTGFPKDLIQVHVEYKKEELQVPTDETDFVCFDGR 128

QY 801 GATTATGCTATATGCTCTCTCTTTATTCGATGCTTGGCAGCTGCGCGGACATAC 860

```

Db 129 DDFHNVVEELLGFLYNSAATDSEKAVEKTLQDMKNPESKEREPEPEPEANSEES 188
QY 861 AGGACATATGATGCCGGGAGGCGCCAGTGGTGGTCAGCAAGTGTTCGAACCGCTGAAC 920
Db 189 DSVFSENTEDLQEQFTTKHSHANSOANHAQEQASFSFEMQLQDKLKVPESENKTS 248
QY 921 AGCTGGCCGATACGTTAATCTGGGATAAGCCGCTGATGCTGTGGCCAGTCAAGAAA 980
Db 249 NSQVSNQODKIDAYKLLKKEMTLDLTKFGSTADALVSDDETRTLVTSLEDDFDELOT 308
QY 981 AAATACGGCTCCAGTATTTCTCGGGAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1040
Db 309 EYAVGCKEDENQEDFELPLFTTDEDMKTPAKSGVEKYPDKQNSNEEDKVLTPV 368
QY 1041 GAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAACATCACCTGGGATCG 1100
Db 369 PGKNDKNTLTWTGDTIFSVTGGEBETRTDMDLESSSEERKEDDDALVDFSGKRPQ 428
QY 1101 GACCTGTATACGGGCGCTGACACAGCCCTCTTGACAGTAACCGGATAAACGGG 1160
Db 429 SATDYSDDPNDVGLFVIDIPKTNNDKVNABHHKGRGVQESKRGVLQVDETELEDEN 488
QY 1161 AAACCTT 1166
Db 489 QEGFKT 494

RESULT 14

US-09-924-400-147/c

; Sequence 147, Application US/09924400

; Patent No. US20020165371A1

; GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.

; APPLICANT: Reed, Steven G.

; APPLICANT: Smith, John M.

; APPLICANT: Misher, Lynda E.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Day, Craig H.

; APPLICANT: Li, Samuel X.

; APPLICANT: Deng, Ta

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.419C12

; CURRENT APPLICATION NUMBER: US/09/924,400

; NUMBER OF SEQ ID NOS: 340

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 147

; LENGTH: 579

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 383, 453, 465, 501

; OTHER INFORMATION: n = A,T,C or G

US-09-924-400-147

Query Match

Best Local Similarity 2.3%; Score 33.8; DB 9; Length 579;

Matches .59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1316 CGCGGGGCGGCATACATTTGGTAAAGCACATGTCGAGCTCTCCTGAAAATGTTGG 1375
Db 427 CGGTGGAGGCGCAGTGGATGGCAGCACACGCGGCGAGCTGGTCNTGTCAGACGGCTTTG 368

QY 1376 CGCAGGCGCTGATGTCACCTGTGGAGGAGCAGGACTGGG 1417

Db 367 CTGACATCATCACCGCTGACACGACGAGGAGTGGAGACGCTGG 326

RESULT 15

US-09-810-936-147/c

; Sequence 147, Application US/09810936

; Patent No. US20020068285A1

; GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.

; APPLICANT: Reed, Steven G.

; APPLICANT: Smith, John M.

; APPLICANT: Misher, Lynda E.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Day, Craig H.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; FILE REFERENCE: 210121.419C11

; CURRENT APPLICATION NUMBER: US/09/810,936

; NUMBER OF SEQ ID NOS: 334

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 147

; LENGTH: 579

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(579)

; OTHER INFORMATION: n = A,T,C or G

US-09-810-936-147

Query Match

Best Local Similarity 2.3%; Score 33.8; DB 10; Length 579;

Matches .59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1316 CGCGGGGCGGCATACATTTGGTAAAGCACATGTCGAGCTCTCCTGAAAATGTTGG 1375
Db 427 CGGTGGAGGCGCAGTGGATGGCAGCACACGCGGCGAGCTGGTCNTGTCAGACGGCTTTG 368

QY 1376 CGCAGGCGCTGATGTCACCTGTGGAGGAGCAGGACTGGG 1417

Db 367 CTGACATCATCACCGCTGACACGACGAGGAGTGGAGACGCTGG 326

RESULT 16

US-09-429-755-147/c

; Sequence 147, Application US/09429755A

; Patent No. US20020111467A1

; GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.

; APPLICANT: Reed, Steven G.

; APPLICANT: Misher, Lynda

; APPLICANT: Retter, Marc W.

; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; FILE REFERENCE: 210121.419C6

; CURRENT APPLICATION NUMBER: US/09/429,755A

; NUMBER OF SEQ ID NOS: 315

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 147

; LENGTH: 579

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(579)

; OTHER INFORMATION: n = A,T,C or G

US-09-429-755-147

Db 426 ILEKEGVKLTPEGEIFPEYDILHKLVSQGIKGGDERGLLSLAFHPNKKNGLYVY 485
Qy 439 CTGTAGTATCTATCCAGCAATAGTATATCTCTGTTCATCAATAAAGTTGACTTTGTA 498
Db 486 TTQERWAGPHHILRVVYTSRKNPHQVDLRTARVLEVAELHRKHLGQLFGPDG 545
Qy 499 TACAACATGGAATTCCTTAAATCCGAGCTATTCGTATGATAAATAAACTCTCTG 558
Db 546 FLYIILGDGMITLDDMEEDGLSDFGVLRLDVTDMCNVYSIPRNPENSTNQPE 605
Qy 559 TCTGATCTCTCGGCGCTATCGGGAGC 587
Db 606 VFAHGLHDPGRCAVDRHPTDININLTLC 634

RESULT 19

US-10-092-154-1415/c
; Sequence 1415, Application US/10092154
; Publication No. US20030054375A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1415
; LENGTH: 6383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1415

Query Match 2.2%; Score 33.4; DB 9; Length 6383;
Best Local Similarity 55.7%; Pred. No. 16;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGCGGAGGCATACATTTGTAAGCACATGTCAGCGTCTCTG 1362
Db 325 CTGCAGGTGTACGGCGTGGAGGCCAGTGAGATGCACACACGCGGCGAGCTGCTG 266
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGCACCTGTGGAGGACGAGGACTGG 1417
Db 265 CAGAACGGCTTTGCTGACATCATCACCCTGTACCAGCAGAAGGTGGAGATGTGG 211

RESULT 20

US-09-764-847-1415/c
; Sequence 1415, Application US/09764847
; Patent No. US20020132767A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1415
; LENGTH: 6383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1415

Query Match 2.2%; Score 33.4; DB 10; Length 6383;
Best Local Similarity 55.7%; Pred. No. 16;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGCGGAGGCATACATTTGTAAGCACATGTCAGCGTCTCTG 1362
Db 325 CTGCAGGTGTACGGCGTGGAGGCCAGTGAGATGCACACACGCGGCGAGCTGCTG 266

Db 325 CTGCAGGTGTACGGCGTGGAGGCCAGTGAGATGCACACACGCGGCGAGCTGCTG 266
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGCACCTGTGGAGGACGAGGACTGG 1417
Db 265 CAGAACGGCTTTGCTGACATCATCACCCTGTACCAGCAGAAGGTGGAGGATGTGG 211

RESULT 21

US-10-092-154-1414/c
; Sequence 1414, Application US/10092154
; Publication No. US20030054375A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1414
; LENGTH: 21732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1414

Query Match 2.2%; Score 33.4; DB 9; Length 21732;
Best Local Similarity 55.7%; Pred. No. 37;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGCGGAGGCATACATTTGTAAGCACATGTCAGCGTCTCTG 1362
Db 15674 CTGCAGGTGTACGGCGTGGAGGCCAGTGAGATGCACACACGCGGCGAGCTGCTG 15615
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGCACCTGTGGAGGACGAGGACTGG 1417
Db 15614 CAGAACGGCTTTGCTGACATCATCACCCTGTACCAGCAGAAGGTGGAGGATGTGG 15560

RESULT 22

US-09-764-847-1414/c
; Sequence 1414, Application US/09764847
; Patent No. US20020132767A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1414
; LENGTH: 21732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1414

Query Match 2.2%; Score 33.4; DB 10; Length 21732;
Best Local Similarity 55.7%; Pred. No. 37;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 1303 CTGTGGCCCTGATCGCGGAGGCATACATTTGTAAGCACATGTCAGCGTCTCTG 1362
Db 15674 CTGCAGGTGTACGGCGTGGAGGCCAGTGAGATGCACACACGCGGCGAGCTGCTG 15615
Qy 1363 AAAAATGATTTGGCGAGGCGCTGATGTGTGCACCTGTGGAGGACGAGGACTGG 1417
Db 15614 CAGAACGGCTTTGCTGACATCATCACCCTGTACCAGCAGAAGGTGGAGGATGTGG 15560

RESULT 23

US-09-918-995-29551

```
; Sequence 29551, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29551
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(484)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-29551

Query Match          2.28; Score 33.2; DB 9; Length 484;
Best Local Similarity 57.88; Pred. No. 2.9; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 43;

QY 1316 CGCGGGAGGCATACATTTGGTAAAGCACATGTCGACGCTCTCCTGAAAAATGTAATTGG 1375
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 149 CGGTGGAGGCCAGTGAGATGGCAGCACACGGGGCAGCTGCTCTCAGAACGCGCTTG 208
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1376 CGCAGGCCCTGTGGTGCACCTGTGGAGGACGAGGACTGGG 1417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 209 CTGACATCATCACCGTGTACCAGCAGAAAGTGGAGGATGTGG 250

RESULT 24
US-09-992-598-352
; Sequence 352, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
```

:	PRIOR APPLICATION NUMBER:	60/089512
:	PRIOR FILING DATE:	1998-06-16
:	PRIOR APPLICATION NUMBER:	60/089514
:	PRIOR FILING DATE:	1998-06-16
:	PRIOR APPLICATION NUMBER:	60/089532
:	PRIOR FILING DATE:	1998-06-17
:	PRIOR APPLICATION NUMBER:	60/089538
:	PRIOR FILING DATE:	1998-06-17
:	PRIOR APPLICATION NUMBER:	60/089598
:	PRIOR FILING DATE:	1998-06-17
:	PRIOR APPLICATION NUMBER:	60/089599
:	PRIOR FILING DATE:	1998-06-17
:	PRIOR APPLICATION NUMBER:	60/089600
:	PRIOR FILING DATE:	1998-06-17
:	PRIOR APPLICATION NUMBER:	60/089653
:	PRIOR FILING DATE:	1998-06-17
:	PRIOR APPLICATION NUMBER:	60/089801
:	PRIOR FILING DATE:	1998-06-18
:	PRIOR APPLICATION NUMBER:	60/089900
:	PRIOR FILING DATE:	1998-06-18
:	PRIOR APPLICATION NUMBER:	60/089908
:	PRIOR FILING DATE:	1998-06-18
:	PRIOR APPLICATION NUMBER:	60/089947
:	PRIOR FILING DATE:	1998-06-19
:	PRIOR APPLICATION NUMBER:	60/089948
:	PRIOR FILING DATE:	1998-06-19
:	PRIOR APPLICATION NUMBER:	60/089952
:	PRIOR FILING DATE:	1998-06-19
:	PRIOR APPLICATION NUMBER:	60/090246
:	PRIOR FILING DATE:	1998-06-22
:	PRIOR APPLICATION NUMBER:	60/090252
:	PRIOR FILING DATE:	1998-06-22
:	PRIOR APPLICATION NUMBER:	60/090254
:	PRIOR FILING DATE:	1998-06-22
:	PRIOR APPLICATION NUMBER:	60/090349
:	PRIOR FILING DATE:	1998-06-23
:	PRIOR APPLICATION NUMBER:	60/090355
:	PRIOR FILING DATE:	1998-06-23
:	PRIOR APPLICATION NUMBER:	60/090429
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090431
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090435
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090444
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090445
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090472
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090535
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090540
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090542
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090557
:	PRIOR FILING DATE:	1998-06-24
:	PRIOR APPLICATION NUMBER:	60/090676
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090678
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090690
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090694
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090695
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090696
:	PRIOR FILING DATE:	1998-06-25
:	PRIOR APPLICATION NUMBER:	60/090862
:	PRIOR FILING DATE:	1998-06-26
:	PRIOR APPLICATION NUMBER:	60/090863

1	PRIOR FILING DATE:	1998-06-26
2	PRIOR APPLICATION NUMBER:	60/091360
3	PRIOR FILING DATE:	1998-07-01
4	PRIOR APPLICATION NUMBER:	60/091478
5	PRIOR FILING DATE:	1998-07-02
6	PRIOR APPLICATION NUMBER:	60/091544
7	PRIOR FILING DATE:	1998-07-01
8	PRIOR APPLICATION NUMBER:	60/091519
9	PRIOR FILING DATE:	1998-07-02
10	PRIOR APPLICATION NUMBER:	60/091626
11	PRIOR FILING DATE:	1998-07-02
12	PRIOR APPLICATION NUMBER:	60/091633
13	PRIOR FILING DATE:	1998-07-02
14	PRIOR APPLICATION NUMBER:	60/091978
15	PRIOR FILING DATE:	1998-07-07
16	PRIOR APPLICATION NUMBER:	60/091982
17	PRIOR FILING DATE:	1998-07-07
18	PRIOR APPLICATION NUMBER:	60/092182
19	PRIOR FILING DATE:	1998-07-09

Query Match	2.2%;	Score 33.2;	DB 9;	Length 3236;
Best Local Similarity	54.0%;	Pred. No. 11;		
Matches 68;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;
Y	1355	GTCTCCTGAAAAATGATTATTTGGCGAGGCGCTGTATGGTGACACTGTGGAGGACGAGGACT	1414	
b	1998	GTCTCATTAAACAATGCAATTTTCAGCTCTGTAGCATTTGGGAAGCTGTCCATTGAAAAGSCCT	2057	
Y	1415	GGGATTTGAAAAATAAATGTGTGTACAGGAACGGCAAAATATACCATCATCACCAGTGGCCCTGGA	1474	
b	2058	TGGATTATTCCTGTACTTTGAACACTGAAATGAAATTTATGCCCGTGTTTTCAAGGTTTGA	2117	
Y	1475	AGGAGC	1480	
	2118	ATGAGC	2123	

RESULT 25
US-09-989-293A-352
Sequence 352, Application US/09989293A
Patent No. US2002017716A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan I.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secured and Trans-
Acids Encoding th
FILE REFERENCE: P2730PIC66
CURRENT APPLICATION NUMBER: US/09/989,
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16

/	PRIOR FILING DATE:	1998-06-16
/	PRIOR APPLICATION NUMBER:	60/089512
/	PRIOR FILING DATE:	1998-06-16
/	PRIOR APPLICATION NUMBER:	60/089514
/	PRIOR FILING DATE:	1998-06-16
/	PRIOR APPLICATION NUMBER:	60/089532
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089538
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089598
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089599
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089600
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089653
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089801
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089907
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089908
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089947
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/089948
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/089952
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/090246
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090252
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090254
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090349
/	PRIOR FILING DATE:	1998-06-23
/	PRIOR APPLICATION NUMBER:	60/090355
/	PRIOR FILING DATE:	1998-06-23
/	PRIOR APPLICATION NUMBER:	60/090429
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090431
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090435
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090444
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090445
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090472
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090535
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090540
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090542
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090557
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090676
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090678
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090690
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090694
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090695
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090696
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090862
/	PRIOR FILING DATE:	1998-06-26

```

; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCGAAATGTTGGCGCAGGCGCTGATGTCACCTGTGGAGGACGGGACT 1414
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1998 GTCCTATTACATGCAATTCAGCTCGTCGAGCATTTGGGAAGCTGTCCATTGAAAAGGCT 2057
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1415 GGGATGGAAAAATAATGTGTGTACAGGAAAGCGCAATATACCATCACCAGTGGCCTGGA 1474
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2058 TGGATTATCCCTGACTTGTGAACATGAACATGAATGATGCCCGTGTTCAGGTTTGA 2117
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 26
US-09-989-735-352
; Sequence 352, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C61
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
```


1	PRIOR APPLICATION NUMBER: 60/049787
2	PRIOR FILING DATE: 1997-06-16
3	PRIOR APPLICATION NUMBER: 60/063250
4	PRIOR FILING DATE: 1997-10-17
5	PRIOR APPLICATION NUMBER: 60/065186
6	PRIOR FILING DATE: 1997-11-12
7	PRIOR APPLICATION NUMBER: 60/065311
8	PRIOR FILING DATE: 1997-11-13
9	PRIOR APPLICATION NUMBER: 60/066770
10	PRIOR FILING DATE: 1997-11-24
11	PRIOR APPLICATION NUMBER: 60/075945
12	PRIOR FILING DATE: 1998-02-25
13	PRIOR APPLICATION NUMBER: 60/078910
14	PRIOR FILING DATE: 1998-03-20
15	PRIOR APPLICATION NUMBER: 60/083322
16	PRIOR FILING DATE: 1998-04-28
17	PRIOR APPLICATION NUMBER: 60/084600
18	PRIOR FILING DATE: 1998-05-07
19	PRIOR APPLICATION NUMBER: 60/087100
20	PRIOR FILING DATE: 1998-05-28
21	PRIOR APPLICATION NUMBER: 60/087607
22	PRIOR FILING DATE: 1998-06-02
23	PRIOR APPLICATION NUMBER: 60/087609
24	PRIOR FILING DATE: 1998-06-02
25	PRIOR APPLICATION NUMBER: 60/087759
26	PRIOR FILING DATE: 1998-06-02
27	PRIOR APPLICATION NUMBER: 60/087827
28	PRIOR FILING DATE: 1998-06-03
29	PRIOR APPLICATION NUMBER: 60/088021
30	PRIOR FILING DATE: 1998-06-04
31	PRIOR APPLICATION NUMBER: 60/088025
32	PRIOR FILING DATE: 1998-06-04
33	PRIOR APPLICATION NUMBER: 60/088026
34	PRIOR FILING DATE: 1998-06-04
35	PRIOR APPLICATION NUMBER: 60/088028
36	PRIOR FILING DATE: 1998-06-04
37	PRIOR APPLICATION NUMBER: 60/088029
38	PRIOR FILING DATE: 1998-06-04
39	PRIOR APPLICATION NUMBER: 60/088030
40	PRIOR FILING DATE: 1998-06-04
41	PRIOR APPLICATION NUMBER: 60/088033
42	PRIOR FILING DATE: 1998-06-04
43	PRIOR APPLICATION NUMBER: 60/088326
44	PRIOR FILING DATE: 1998-06-04
45	PRIOR APPLICATION NUMBER: 60/088167
46	PRIOR FILING DATE: 1998-06-05
47	PRIOR APPLICATION NUMBER: 60/088202
48	PRIOR FILING DATE: 1998-06-05
49	PRIOR APPLICATION NUMBER: 60/088212
50	PRIOR FILING DATE: 1998-06-05
51	PRIOR APPLICATION NUMBER: 60/088217
52	PRIOR FILING DATE: 1998-06-05
53	PRIOR APPLICATION NUMBER: 60/088655
54	PRIOR FILING DATE: 1998-06-09
55	PRIOR APPLICATION NUMBER: 60/088734
56	PRIOR FILING DATE: 1998-06-10
57	PRIOR APPLICATION NUMBER: 60/088738
58	PRIOR FILING DATE: 1998-06-10
59	PRIOR APPLICATION NUMBER: 60/088742
60	PRIOR FILING DATE: 1998-06-10
61	PRIOR APPLICATION NUMBER: 60/088810
62	PRIOR FILING DATE: 1998-06-10
63	PRIOR APPLICATION NUMBER: 60/088824
64	PRIOR FILING DATE: 1998-06-10
65	PRIOR APPLICATION NUMBER: 60/088826
66	PRIOR FILING DATE: 1998-06-10
67	PRIOR APPLICATION NUMBER: 60/088858
68	PRIOR FILING DATE: 1998-06-11
69	PRIOR APPLICATION NUMBER: 60/088861
70	PRIOR FILING DATE: 1998-06-11
71	PRIOR APPLICATION NUMBER: 60/088876
72	PRIOR FILING DATE: 1998-06-11
73	PRIOR APPLICATION NUMBER: 60/089105

1	PRIOR FILING DATE: 1998-06-12
2	PRIOR APPLICATION NUMBER: 60/089440
3	PRIOR FILING DATE: 1998-06-16
4	PRIOR APPLICATION NUMBER: 60/089512
5	PRIOR FILING DATE: 1998-06-16
6	PRIOR APPLICATION NUMBER: 60/089514
7	PRIOR FILING DATE: 1998-06-16
8	PRIOR APPLICATION NUMBER: 60/089532
9	PRIOR FILING DATE: 1998-06-17
10	PRIOR APPLICATION NUMBER: 60/089538
11	PRIOR FILING DATE: 1998-06-17
12	PRIOR APPLICATION NUMBER: 60/089598
13	PRIOR FILING DATE: 1998-06-17
14	PRIOR APPLICATION NUMBER: 60/089599
15	PRIOR FILING DATE: 1998-06-17
16	PRIOR APPLICATION NUMBER: 60/089600
17	PRIOR FILING DATE: 1998-06-17
18	PRIOR APPLICATION NUMBER: 60/089653
19	PRIOR FILING DATE: 1998-06-17
20	PRIOR APPLICATION NUMBER: 60/089801
21	PRIOR FILING DATE: 1998-06-18
22	PRIOR APPLICATION NUMBER: 60/089907
23	PRIOR FILING DATE: 1998-06-18
24	PRIOR APPLICATION NUMBER: 60/089908
25	PRIOR FILING DATE: 1998-06-18
26	PRIOR APPLICATION NUMBER: 60/089947
27	PRIOR FILING DATE: 1998-06-19
28	PRIOR APPLICATION NUMBER: 60/089948
29	PRIOR FILING DATE: 1998-06-19
30	PRIOR APPLICATION NUMBER: 60/089952
31	PRIOR FILING DATE: 1998-06-19
32	PRIOR APPLICATION NUMBER: 60/090246
33	PRIOR FILING DATE: 1998-06-22
34	PRIOR APPLICATION NUMBER: 60/090252
35	PRIOR FILING DATE: 1998-06-22
36	PRIOR APPLICATION NUMBER: 60/090254
37	PRIOR FILING DATE: 1998-06-22
38	PRIOR APPLICATION NUMBER: 60/090349
39	PRIOR FILING DATE: 1998-06-23
40	PRIOR APPLICATION NUMBER: 60/090355
41	PRIOR FILING DATE: 1998-06-23
42	PRIOR APPLICATION NUMBER: 60/090429
43	PRIOR FILING DATE: 1998-06-24
44	PRIOR APPLICATION NUMBER: 60/090431
45	PRIOR FILING DATE: 1998-06-24
46	PRIOR APPLICATION NUMBER: 60/090435
47	PRIOR FILING DATE: 1998-06-24
48	PRIOR APPLICATION NUMBER: 60/090444
49	PRIOR FILING DATE: 1998-06-24
50	PRIOR APPLICATION NUMBER: 60/090445
51	PRIOR FILING DATE: 1998-06-24
52	PRIOR APPLICATION NUMBER: 60/090472
53	PRIOR FILING DATE: 1998-06-24
54	PRIOR APPLICATION NUMBER: 60/090535
55	PRIOR FILING DATE: 1998-06-24
56	PRIOR APPLICATION NUMBER: 60/090540
57	PRIOR FILING DATE: 1998-06-24
58	PRIOR APPLICATION NUMBER: 60/090542
59	PRIOR FILING DATE: 1998-06-24
60	PRIOR APPLICATION NUMBER: 60/090590
61	PRIOR FILING DATE: 1998-06-25
62	PRIOR APPLICATION NUMBER: 60/090694
63	PRIOR FILING DATE: 1998-06-25
64	PRIOR APPLICATION NUMBER: 60/090695
65	PRIOR FILING DATE: 1998-06-25
66	PRIOR APPLICATION NUMBER: 60/090696
67	PRIOR FILING DATE: 1998-06-25

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	8												

1	PRIOR FILING DATE: 1998-06-11
2	PRIOR APPLICATION NUMBER: 60/089105
3	PRIOR FILING DATE: 1998-06-12
4	PRIOR APPLICATION NUMBER: 60/089440
5	PRIOR FILING DATE: 1998-06-16
6	PRIOR APPLICATION NUMBER: 60/089512
7	PRIOR FILING DATE: 1998-06-16
8	PRIOR APPLICATION NUMBER: 60/089514
9	PRIOR FILING DATE: 1998-06-16
10	PRIOR APPLICATION NUMBER: 60/089532
11	PRIOR FILING DATE: 1998-06-17
12	PRIOR APPLICATION NUMBER: 60/089538
13	PRIOR FILING DATE: 1998-06-17
14	PRIOR APPLICATION NUMBER: 60/089598
15	PRIOR FILING DATE: 1998-06-17
16	PRIOR APPLICATION NUMBER: 60/089599
17	PRIOR FILING DATE: 1998-06-17
18	PRIOR APPLICATION NUMBER: 60/089600
19	PRIOR FILING DATE: 1998-06-17
20	PRIOR APPLICATION NUMBER: 60/089653
21	PRIOR FILING DATE: 1998-06-17
22	PRIOR APPLICATION NUMBER: 60/089801
23	PRIOR FILING DATE: 1998-06-18
24	PRIOR APPLICATION NUMBER: 60/089907
25	PRIOR FILING DATE: 1998-06-18
26	PRIOR APPLICATION NUMBER: 60/089908
27	PRIOR FILING DATE: 1998-06-18
28	PRIOR APPLICATION NUMBER: 60/089947
29	PRIOR FILING DATE: 1998-06-19
30	PRIOR APPLICATION NUMBER: 60/089948
31	PRIOR FILING DATE: 1998-06-19
32	PRIOR APPLICATION NUMBER: 60/089952
33	PRIOR FILING DATE: 1998-06-19
34	PRIOR APPLICATION NUMBER: 60/090246
35	PRIOR FILING DATE: 1998-06-22
36	PRIOR APPLICATION NUMBER: 60/090252
37	PRIOR FILING DATE: 1998-06-22
38	PRIOR APPLICATION NUMBER: 60/090254
39	PRIOR FILING DATE: 1998-06-22
40	PRIOR APPLICATION NUMBER: 60/090349
41	PRIOR FILING DATE: 1998-06-23
42	PRIOR APPLICATION NUMBER: 60/090355
43	PRIOR FILING DATE: 1998-06-23
44	PRIOR APPLICATION NUMBER: 60/090429
45	PRIOR FILING DATE: 1998-06-24
46	PRIOR APPLICATION NUMBER: 60/090431
47	PRIOR FILING DATE: 1998-06-24
48	PRIOR APPLICATION NUMBER: 60/090435
49	PRIOR FILING DATE: 1998-06-24
50	PRIOR APPLICATION NUMBER: 60/090444
51	PRIOR FILING DATE: 1998-06-24
52	PRIOR APPLICATION NUMBER: 60/090445
53	PRIOR FILING DATE: 1998-06-24
54	PRIOR APPLICATION NUMBER: 60/090472
55	PRIOR FILING DATE: 1998-06-24
56	PRIOR APPLICATION NUMBER: 60/090535
57	PRIOR FILING DATE: 1998-06-24
58	PRIOR APPLICATION NUMBER: 60/090540
59	PRIOR FILING DATE: 1998-06-24
60	PRIOR APPLICATION NUMBER: 60/090542
61	PRIOR FILING DATE: 1998-06-24
62	PRIOR APPLICATION NUMBER: 60/090557
63	PRIOR FILING DATE: 1998-06-24
64	PRIOR APPLICATION NUMBER: 60/090676
65	PRIOR FILING DATE: 1998-06-25
66	PRIOR APPLICATION NUMBER: 60/090678
67	PRIOR FILING DATE: 1998-06-25
68	PRIOR APPLICATION NUMBER: 60/090690
69	PRIOR FILING DATE: 1998-06-25
70	PRIOR APPLICATION NUMBER: 60/090694
71	PRIOR FILING DATE: 1998-06-25
72	PRIOR APPLICATION NUMBER: 60/090695
73	PRIOR FILING DATE: 1998-06-25


```

; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTGTAATAATGTTGGCGCAGGCGCTGATGTCGACCTGTGGAGGACGGGACT 1414
Db 1998 GTCTATTAAACATGCAATTCAGTCGTCAGCATTTGGAGCTGCCATTGAAAAGGCT 2057

QY 1415 GGGATGGAATAATAATGTGGTACAGGAACGGCAATATACCATCACCAGTGGCCTGGA 1474
Db 2058 TGGATTATTCCTGACTTGTAAACATGAACATGAATATGCCCCGTGTTCAAGGTTGA 2117

QY 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 30
US-09-991-181-352
; Sequence 352, Application Us/09991181
; Publication No. US20020197615A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730PIC53
; CURRENT APPLICATION NUMBER: US/09/991,181
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
```

; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.0%; Pred. No. 11;

Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCTCTGAAATGTTATGGCCAGGCGCTGTGGTGGACCTGATGGTGGACCTGTGGAGGAGGAGGACT 1414

Db 1998 GTCTCATTAACAATGCAATTCAGCTCTGTCAGCATTTGGGAAGCTGTCCATTGAAAGGCCT 2057

QY 1415 GGGATGGAAAAATAATGTGTGTACAGGAACGGCAAAATATACCATCACCAGTGGCCTGGA 1474

Db 2058 TGGATTTATCCCTGTACTTGAACATCAAACTGAATTAATGCCGCTGTTCAAGTTGA 2117

QY 1475 AGGAGC 1480

Db 2118 ATGAGC 2123

RESULT 31

US-09-993-687-352

; Sequence 352, Application US/09993687

; Publication No. US20020198149A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE OF INVENTION: ACIDS Encoding the Same
FILE REFERENCE: P273PICI11
CURRENT APPLICATION NUMBER: US/09/393,687
CURRENT FILING DATE: 2002-11-14
PRIOR APPLICATION NUMBER: 60/045787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861

; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCTCTGAAATATGTTGGCCAGGCGCTGTGTCACCTGTGGAGGAGCAGGACT 1414
Db 1998 GTCTCATTACATGTTTCACCTGTGGAAGCTGTCCATTGAAAGGCCT 2057
QY 1415 GGGATGGAATAAATATGTTACAGGAACGCAATATACCATCACAGTGCCTGGA 1474
Db 2058 TGGATTATCCCTGTACTTGAACATGAACACTGAATATTGCCCGTTCAGAGTTTGA 2117
QY 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 32
US-09-989-734-352
; Sequence 352, Application US/09989734
; Publication No. US20030003531A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerriksen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC64
; CURRENT APPLICATION NUMBER: US/09/989,734
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11

```

; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694

Query Match      2.28; Score 33.2; DB 9; Length 3236;
Best Local Similarity 54.08; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCTCTGAAAAATGTTATGGCGCGGCTGTGTGTCACCTGTGTGAGGAGCAGGACT 1414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1998 GTCTCATTAACATGTCATTCAGCTCTCAGCATGGGAGCTGTCCATTTGAAAAGGCCT 2057

QY 1415 GGGATGGAAAAATAATGTGTGTACAGGAAACGGCAATATATACCATCACCAGTGGCTGGA 1474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2058 TGGATTTATCCCTGTACTTGAACATGAAACTGAATTTATGCCCGTGTTCAGGTTTGA 2117

QY 1475 AGGAGC 1480
    |||||
Db 2118 ATGAGC 2123

RESULT 33
US-10-028-072-463
; Sequence 463, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; FILE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
```

;	PRIOR APPLICATION NUMBER:	60/056974
;	PRIOR FILING DATE:	1997-08-26
;	PRIOR APPLICATION NUMBER:	60/059113
;	PRIOR FILING DATE:	1997-09-17
;	PRIOR APPLICATION NUMBER:	60/059115
;	PRIOR FILING DATE:	1997-09-17
;	PRIOR APPLICATION NUMBER:	60/059117
;	PRIOR FILING DATE:	1997-09-17
;	PRIOR APPLICATION NUMBER:	60/059122
;	PRIOR FILING DATE:	1997-09-17
;	PRIOR APPLICATION NUMBER:	60/059184
;	PRIOR FILING DATE:	1997-09-17
;	PRIOR APPLICATION NUMBER:	60/059263
;	PRIOR FILING DATE:	1997-09-18
;	PRIOR APPLICATION NUMBER:	60/059352
;	PRIOR FILING DATE:	1997-09-19
;	PRIOR APPLICATION NUMBER:	60/059588
;	PRIOR FILING DATE:	1997-09-19
;	PRIOR APPLICATION NUMBER:	60/059835
;	PRIOR FILING DATE:	1997-09-24
;	PRIOR APPLICATION NUMBER:	60/062250
;	PRIOR FILING DATE:	1997-10-17
;	PRIOR APPLICATION NUMBER:	60/062285
;	PRIOR FILING DATE:	1997-10-17
;	PRIOR APPLICATION NUMBER:	60/062287
;	PRIOR FILING DATE:	1997-10-17
;	PRIOR APPLICATION NUMBER:	60/062814
;	PRIOR FILING DATE:	1997-10-24
;	PRIOR APPLICATION NUMBER:	60/062816
;	PRIOR FILING DATE:	1997-10-24
;	PRIOR APPLICATION NUMBER:	60/063045
;	PRIOR FILING DATE:	1997-10-24
;	PRIOR APPLICATION NUMBER:	60/063082
;	PRIOR FILING DATE:	1997-10-31
;	PRIOR APPLICATION NUMBER:	60/063127
;	PRIOR FILING DATE:	1997-10-24
;	PRIOR APPLICATION NUMBER:	60/063327
;	PRIOR FILING DATE:	1997-10-27
;	PRIOR APPLICATION NUMBER:	60/063329
;	PRIOR FILING DATE:	1997-10-27
;	PRIOR APPLICATION NUMBER:	60/063550
;	PRIOR FILING DATE:	1997-10-28
;	PRIOR APPLICATION NUMBER:	60/063561
;	PRIOR FILING DATE:	1997-10-28
;	PRIOR APPLICATION NUMBER:	60/063704
;	PRIOR FILING DATE:	1997-10-29
;	PRIOR APPLICATION NUMBER:	60/063733
;	PRIOR FILING DATE:	1997-10-29
;	PRIOR APPLICATION NUMBER:	60/063755
;	PRIOR FILING DATE:	1997-10-17
;	PRIOR APPLICATION NUMBER:	60/064248
;	PRIOR FILING DATE:	1997-11-03
;	PRIOR APPLICATION NUMBER:	60/064809
;	PRIOR FILING DATE:	1997-11-07
;	PRIOR APPLICATION NUMBER:	60/065186
;	PRIOR FILING DATE:	1997-11-12
;	PRIOR APPLICATION NUMBER:	60/065846
;	PRIOR FILING DATE:	1997-11-17
;	PRIOR APPLICATION NUMBER:	60/066364
;	PRIOR FILING DATE:	1997-11-21
;	PRIOR APPLICATION NUMBER:	60/066453
;	PRIOR FILING DATE:	1997-11-24
;	PRIOR APPLICATION NUMBER:	60/066511
;	PRIOR FILING DATE:	1997-11-24
;	PRIOR APPLICATION NUMBER:	60/066770
;	PRIOR FILING DATE:	1997-11-24
;	PRIOR APPLICATION NUMBER:	60/069212
;	PRIOR FILING DATE:	1997-12-11
;	PRIOR APPLICATION NUMBER:	60/069278

60/069334	PRIOR APPLICATION NUMBER: 60/069334	PRIOR FILING DATE: 1997-12-11
60/069694	PRIOR APPLICATION NUMBER: 60/069694	PRIOR FILING DATE: 1997-12-11
60/072320	PRIOR APPLICATION NUMBER: 60/072320	PRIOR FILING DATE: 1997-12-16
60/073612	PRIOR APPLICATION NUMBER: 60/073612	PRIOR FILING DATE: 1998-01-23
60/074086	PRIOR APPLICATION NUMBER: 60/074086	PRIOR FILING DATE: 1998-02-04
60/074092	PRIOR APPLICATION NUMBER: 60/074092	PRIOR FILING DATE: 1998-02-09
60/077791	PRIOR APPLICATION NUMBER: 60/077791	PRIOR FILING DATE: 1998-02-09
60/078910	PRIOR APPLICATION NUMBER: 60/078910	PRIOR FILING DATE: 1998-03-12
60/079294	PRIOR APPLICATION NUMBER: 60/079294	PRIOR FILING DATE: 1998-03-20
60/079663	PRIOR APPLICATION NUMBER: 60/079663	PRIOR FILING DATE: 1998-03-25
60/079728	PRIOR APPLICATION NUMBER: 60/079728	PRIOR FILING DATE: 1998-02-27
60/080165	PRIOR APPLICATION NUMBER: 60/080165	PRIOR FILING DATE: 1998-03-27
60/081203	PRIOR APPLICATION NUMBER: 60/081203	PRIOR FILING DATE: 1998-03-31
60/081229	PRIOR APPLICATION NUMBER: 60/081229	PRIOR FILING DATE: 1998-04-09
60/081695	PRIOR APPLICATION NUMBER: 60/081695	PRIOR FILING DATE: 1998-04-09
60/081818	PRIOR APPLICATION NUMBER: 60/081818	PRIOR FILING DATE: 1998-04-15
60/082999	PRIOR APPLICATION NUMBER: 60/082999	PRIOR FILING DATE: 1998-04-15
60/083322	PRIOR APPLICATION NUMBER: 60/083322	PRIOR FILING DATE: 1998-04-24
60/083545	PRIOR APPLICATION NUMBER: 60/083545	PRIOR FILING DATE: 1998-04-28
60/084637	PRIOR APPLICATION NUMBER: 60/084637	PRIOR FILING DATE: 1998-04-29
60/085149	PRIOR APPLICATION NUMBER: 60/085149	PRIOR FILING DATE: 1998-05-07
60/085323	PRIOR APPLICATION NUMBER: 60/085323	PRIOR FILING DATE: 1998-05-12
60/085338	PRIOR APPLICATION NUMBER: 60/085338	PRIOR FILING DATE: 1998-05-13
60/085339	PRIOR APPLICATION NUMBER: 60/085339	PRIOR FILING DATE: 1998-05-13
60/085579	PRIOR APPLICATION NUMBER: 60/085579	PRIOR FILING DATE: 1998-05-15
60/085697	PRIOR APPLICATION NUMBER: 60/085697	PRIOR FILING DATE: 1998-05-15
60/085704	PRIOR APPLICATION NUMBER: 60/085704	PRIOR FILING DATE: 1998-05-15
60/086414	PRIOR APPLICATION NUMBER: 60/086414	PRIOR FILING DATE: 1998-05-22
60/086430	PRIOR APPLICATION NUMBER: 60/086430	PRIOR FILING DATE: 1998-05-22
60/087106	PRIOR APPLICATION NUMBER: 60/087106	PRIOR FILING DATE: 1998-05-28
60/088026	PRIOR APPLICATION NUMBER: 60/088026	PRIOR FILING DATE: 1998-06-04
60/088730	PRIOR APPLICATION NUMBER: 60/088730	PRIOR FILING DATE: 1998-06-10
60/088741	PRIOR APPLICATION NUMBER: 60/088741	PRIOR FILING DATE: 1998-06-10

;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 19/98-06-11
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090538
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

Query Match 2.2%; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.0%; Pred No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY	1355	GTCTCTGAAATATGTTGGCGCAGGCGCTGATGTCACCTGTGGAGGAGGAGGACT	1414
Db	1998		
QY	1415	GGATGGGAAATAATGTGTACAGGAACGGCAATATACCATCACCCTGGCGTGA	1474
Db	2058		
QY	1475	TGATTTATCCCTGTACTTGAACATGAATGAAATATGCTGCTTTCAGGTTTGA	2117
Db	2118	ATGAGC 2123	

RESULT 34

US-09-997-653-352
; Sequence 352, Application US/09997653
; Publication No. US2003008297A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey

;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2730PIC38
;; CURRENT APPLICATION NUMBER: US/09/997,653
;; CURRENT FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690

PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 2.28; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.08; Pred. No. 11;

Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 1355 GTCTCTGAAATAATGTTGGCGCAGGCGCTGATGGTGACCTGTGGAGGAGCAGGACT 1414

Db 1998 GTCTCATTAACATGCAATTCAGCTGTCACCATTTGGAGCTGTCCATTGAAAGGCT 2057

Oy 1415 GGGATGAAAAATAAATGTGTACAGGAAACGCAATATACCATCACCAGTGGCCTGGA 1474

Db 2058 TGGATTATCCCTGCTACTTGAACATGAAACTGAAATATATGCCCGTGTTCGAAGTTGA 2117

Oy 1475 AGGAGC 1480

Db 2118 ATGAGC 2123

RESULT 35

US-09-993-667-352

; Sequence 352, Application US/09993667

; Publication No. US20030022187A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tamas, Daniel

; APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PlC4
CURRENT APPLICATION NUMBER: US/09/993.667
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

```

RESULT 36
US-10-121-049-463
; Sequence 463, Application US/10121049
; Publication No. US20030022339A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRA
; TITLE OF INVENTION: ACIDS ENCODING
; FILE REFERENCE: P3330R1C17

```

	Query Match	2.2%	Score 33.2;	DB 9;	Length 3226;
	Best Local Similarity	54.0%;	Pred. No. 11;		
	Matches 68;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;
Qy	1355	GTCTCTGAAAAATGTAATGTCGCGCAGGGCCTGTAGTGTGCACCTGTGGAGGAGCAGGGACT	1414		
Db	1998	GTCTCATTAACATGCATTTTCAGCTCGTCAGCATTTGGAAAGCTGTCCATTGAAAAAGCCCT	2057		
Qy	1415	GGGATGGAATAATTAATGTGTGTACAGGAACGGCAATATATACCATACACAGTGGCCCTGGA	1474		
Db	2058	TGGATTTATCCCTCTACTCTGAAACACATGAACATGAAATATATGCCCCGTGCTTTTCAAGGTTTGA	2117		

/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090540
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090542
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090557
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090676
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090678
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090690
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090694
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090695
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090696
/	PRIOR FILING DATE:	1998-06-25
/	PRIOR APPLICATION NUMBER:	60/090862
/	PRIOR FILING DATE:	1998-06-26
/	PRIOR APPLICATION NUMBER:	60/090863
/	PRIOR FILING DATE:	1998-06-26
/	PRIOR APPLICATION NUMBER:	60/091360
/	PRIOR FILING DATE:	1998-07-01
/	PRIOR APPLICATION NUMBER:	60/091478
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091544
/	PRIOR FILING DATE:	1998-07-01
/	PRIOR APPLICATION NUMBER:	60/091519
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091626
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091633
/	PRIOR FILING DATE:	1998-07-02
/	PRIOR APPLICATION NUMBER:	60/091978
/	PRIOR FILING DATE:	1998-07-07
/	PRIOR APPLICATION NUMBER:	60/091982
/	PRIOR FILING DATE:	1998-07-07
/	PRIOR APPLICATION NUMBER:	60/092182
/	PRIOR FILING DATE:	1998-07-09

Query Match	2.2%	Score
Best Local Similarity	54.0%	Predefined
Matches	68;	Conservative
	0;	M

Qy	1355	GTCCTCGTGAATAATGATTTGGCGC
Db	1998	GTCCTCATTAACAATGCATTTTCAGC
Qy	1415	GGGTGGGAAAAATAAATGCTGGTAGC
Db	2098	TGGATTATCCCTCTACTTGAAAC
Qy	1475	AGGAGC 1480
Db	2118	ATGAGC 2123

RESULT 40
US-09-990-562-352
; Sequence 352, Application US/0999056
; Publication No. US20030027985A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.

Query Match	2.2%	Score 33.2;	DB 9;	Length 3226;	
Best Local Similarity	54.0%;	Pred. No. 11;			
Matches	68;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;

QY	1355	GTCTCTGAAAAATGTTGGCCGAGGCCGTGATGTGTGCACCTGTGTGGAGGAGCAGCGACT	1414
Db	1998	GTCTCATTAACAATGCAATTCAGTCTGTCAGCAATGGGAAGCTGTCCATTGAAAAGCCCT	2057
QY	1415	GGGTGTGAAAAATAAATGTGTTACAGGAAACGGCAATAATACCATCACCAGTGCCTGGA	1474
Db	2058	TGGATTTATCCCTGTACTTGAACACATGAAACTGAAATTAATGCCCGTGTTCGAAGGTTGA	2117
QY	1475	AGGAGC	1480
Db	2118	ATGAGC	2123

RESULT 40

US-09-990-562-352

; Sequence 352, Application US/09990562

; Publication No. US20030027985A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Marv E.

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pap, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C18
CURRENT APPLICATION NUMBER: US/09/990,562
CURRENT FILING DATE: 2001-11-14
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24

7	PRIOR APPLICATION NUMBER: 60/090535
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090540
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090542
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090557
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090676
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090678
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090690
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090694
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090695
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090696
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090862
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/090863
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/091360
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091478
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091544
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091519
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091626
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091633
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091978
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/091982
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/092182
7	PRIOR FILING DATE: 1998-07-09

Query Match 2.28; Score 33.2; DB 9; Length 3226;

Query Match	2.2%	Score 33.2	DB 9	Length 3226
Best Local Similarity	54.0%	Pred. No. 11		
Matches 68	Conservative	0	Mismatches 58	Indels 0
Matches 0				Gaps 0

QY 1355 GTCTCCTGAAAAATGTATTGGCGCAGGGCCTGATGGTGCACCTGTGGAGGAGCAGGGACT 1414

Db 1998 GTCTCATTAACAATGCATTTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057

1415 GGGATGGAAAAATAAATGTGGTACAGGAAACGGCAAATATACCATCACCAGTGGCCCTGGA 1474

2058 TGGATTTATCCCTGTACTTGAACACATGAACACTGAATATATGCCCGTGTGTTTCAAGGTTTGA 2117

QY 1475 AGGAGC 1480

db . 2118 ATGAGC 2123

RESULT 41

RESOUR 41
US-09-997-428-352

US 03 357 428 352
; Sequence 352, Application US/09997428

; Publication No. US20030027162A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: **FERRARA, NAPOLEONE**

APPLICANT: Fong, Sherman

APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter

1 APPLICANT: Gerritsen, Mary E.
 2 APPLICANT: Goddard, Audrey
 3 APPLICANT: Godowski, Paul J.
 4 APPLICANT: Grimaldi, J. Christopher
 5 APPLICANT: Gurney, Austin L.
 6 APPLICANT: Kljavin, Ivar J.
 7 APPLICANT: Napier, Mary A.
 8 APPLICANT: Pan, James
 9 APPLICANT: Paoni, Nicholas F.
 10 APPLICANT: Roy, Margaret Ann
 11 APPLICANT: Stewart, Timothy A.
 12 APPLICANT: Tumas, Daniel
 13 APPLICANT: Watanabe, Colin K.
 14 APPLICANT: Williams, P. Mickey
 15 APPLICANT: Wood, William I.
 16 APPLICANT: Zhang, Zemin
 17 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 18 TITLE OF INVENTION: Acids Encoding the Same
 19 FILE REFERENCE: P2730PIC44
 20 CURRENT APPLICATION NUMBER: US/09/997,428
 21 CURRENT FILING DATE: 2001-11-15
 22 PRIOR APPLICATION NUMBER: 60/049787
 23 PRIOR FILING DATE: 1997-06-16
 24 PRIOR APPLICATION NUMBER: 60/062250
 25 PRIOR FILING DATE: 1997-10-17
 26 PRIOR APPLICATION NUMBER: 60/065186
 27 PRIOR FILING DATE: 1997-11-12
 28 PRIOR APPLICATION NUMBER: 60/065311
 29 PRIOR FILING DATE: 1997-11-13
 30 PRIOR APPLICATION NUMBER: 60/066770
 31 PRIOR FILING DATE: 1997-11-24
 32 PRIOR APPLICATION NUMBER: 60/075945
 33 PRIOR FILING DATE: 1998-02-25
 34 PRIOR APPLICATION NUMBER: 60/078910
 35 PRIOR FILING DATE: 1998-03-20
 36 PRIOR APPLICATION NUMBER: 60/083322
 37 PRIOR FILING DATE: 1998-04-28
 38 PRIOR APPLICATION NUMBER: 60/084600
 39 PRIOR FILING DATE: 1998-05-07
 40 PRIOR APPLICATION NUMBER: 60/087106
 41 PRIOR FILING DATE: 1998-05-28
 42 PRIOR APPLICATION NUMBER: 60/087607
 43 PRIOR FILING DATE: 1998-06-02
 44 PRIOR APPLICATION NUMBER: 60/087609
 45 PRIOR FILING DATE: 1998-06-02
 46 PRIOR APPLICATION NUMBER: 60/087759
 47 PRIOR FILING DATE: 1998-06-02
 48 PRIOR APPLICATION NUMBER: 60/087827
 49 PRIOR FILING DATE: 1998-06-03
 50 PRIOR APPLICATION NUMBER: 60/088021
 51 PRIOR FILING DATE: 1998-06-04
 52 PRIOR APPLICATION NUMBER: 60/088025
 53 PRIOR FILING DATE: 1998-06-04
 54 PRIOR APPLICATION NUMBER: 60/088026
 55 PRIOR FILING DATE: 1998-06-04
 56 PRIOR APPLICATION NUMBER: 60/088028
 57 PRIOR FILING DATE: 1998-06-04
 58 PRIOR APPLICATION NUMBER: 60/088029
 59 PRIOR FILING DATE: 1998-06-04
 60 PRIOR APPLICATION NUMBER: 60/088030
 61 PRIOR FILING DATE: 1998-06-04
 62 PRIOR APPLICATION NUMBER: 60/088033
 63 PRIOR FILING DATE: 1998-06-04
 64 PRIOR APPLICATION NUMBER: 60/088326
 65 PRIOR FILING DATE: 1998-06-04
 66 PRIOR APPLICATION NUMBER: 60/088167
 67 PRIOR FILING DATE: 1998-06-05
 68 PRIOR APPLICATION NUMBER: 60/088202
 69 PRIOR FILING DATE: 1998-06-05
 70 PRIOR APPLICATION NUMBER: 60/088212
 71 PRIOR FILING DATE: 1998-06-05
 72 PRIOR APPLICATION NUMBER: 60/088217
 73 PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089512
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089538
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089600
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472

; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
 Best Local Similarity 54.0%; Pred. No. 11;
 Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 1355 GTCTCTGAAAAATGTTATGCGCGAGGCGCTGTGAGTGTGACCTGTGAGGAGCAGGACT 1414
 Db 1998 GTCTCATTAACAATGCATTTTCAGCTCGTCAGCATTTGGGAAGCTGCTCCATTGAAAAAGCCT 2057
 QY 1415 GGGATGAAAAATAAATGTGTACAGGAAACGGCAATATACCATCACCAGTGCCTGGA 1474
 Db 2058 TGGATTATTCCTGCTACTTGAACATGAAATGAAATGTCGCGGTGTTCAAGGTTGA 2117
 QY 1475 AGGAGC 1480
 Db 2118 ATGAGC 2123

RESULT 42
 US-09-997-666-352
 ; Sequence 352, Application US/09997666
 ; Publication No. US20030027163A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC42
CURRENT APPLICATION NUMBER: US/09/997,666
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
 Best Local Similarity 54.0%; Pred. No. 11;
 Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 1355 GTCCTCGAAAAATGATTGGCGCAGGCGCTGATGGTGACCTGTGGAGGAGGAGGACT 1414
 DB 1998 GTCTCATTAACAATGCAATTCAGCTGCTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057
 QY 1415 GGGATGGAATAAATGTTGGTACAGAAACGGCAATATACCATCACCAGTGGCCTGGA 1474
 DB 2058 TGGATTATCCCTGCTACTTGAACATGAAACTGAAATATATGCCCCGTTTCAAGGTTGA 2117
 QY 1475 AGGAGC 1480
 DB 2118 ATGAGC 2123

RESULT 43
 US-10-175-746-463
 ; Sequence 463, Application US/10175746
 ; Publication No. US2003002720A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C353
 ; CURRENT APPLICATION NUMBER: US/10/175,746
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 463
 ; LENGTH: 3226
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-175-746-463
 Query Match 2.2%; Score 33.2; DB 9; Length 3226;
 Best Local Similarity 54.0%; Pred. No. 11;
 Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 1355 GTCCTCGAAAAATGATTGGCGCAGGCGCTGATGGTGACCTGTGGAGGAGGAGGACT 1414
 DB 1998 GTCTCATTAACAATGCAATTCAGCTGCTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057
 QY 1415 GGGATGGAATAAATGTTGGTACAGAAACGGCAATATACCATCACCAGTGGCCTGGA 1474
 DB 2058 TGGATTATCCCTGCTACTTGAACATGAAACTGAAATATATGCCCCGTTTCAAGGTTGA 2117
 QY 1475 AGGAGC 1480
 DB 2118 ATGAGC 2123

RESULT 44
 US-10-176-918-463
 ; Sequence 463, Application US/10176918
 ; Publication No. US2003002725A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C382
 ; CURRENT APPLICATION NUMBER: US/10/176,918
 ; CURRENT FILING DATE: 2002-06-20
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 463
 ; LENGTH: 3226
 ; TYPE: DNA

ORGANISM: Homo Sapien
US-10-176-918-463

Query Match 2.28; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.08; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1355 GTCTCCGAAATGATTGGCGCAGGCGCTGATGGTGACCTGTGGAGGACGAGGACT 1414
Db 1998 GTCTCATTAACAATGATTTCAGCTGCTGACGATGGGAAGCTGTCCATTGAAAGGCT 2057
Qy 1415 GGGATGGAATAAATGTTGTACAGGAAGCGCAATATACCATCACCAGTGCCCTGGA 1474
Db 2058 TGGATTATCCCTGACTTGAACATGAACATGAATATATGCCGCTTTCAAGTTTGA 2117
Qy 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 45

US-10-176-921-463
Sequence 463, Application US/10176921
Publication No. US20030027276A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

Acids Encoding the Same

FILE REFERENCE: P330RIC288

CURRENT APPLICATION NUMBER: US/10/176.921

CURRENT FILING DATE: 2002-06-20

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 463

LENGTH: 3226

TYPE: DNA

ORGANISM: Homo Sapien

US-10-176-921-463

Query Match 2.28; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.08; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1355 GTCTCCGAAATGATTGGCGCAGGCGCTGATGGTGACCTGTGGAGGACGAGGACT 1414
Db 1998 GTCTCATTAACAATGATTTCAGCTGCTGACGATGGGAAGCTGTCCATTGAAAGGCT 2057
Qy 1415 GGGATGGAATAAATGTTGTACAGGAAGCGCAATATACCATCACCAGTGCCCTGGA 1474
Db 2058 TGGATTATCCCTGACTTGAACATGAACATGAATATATGCCGCTTTCAAGTTTGA 2117
Qy 1475 AGGAGC 1480
Db 2118 ATGAGC 2123

RESULT 46

US-09-990-711-352
Sequence 352, Application US/09990711
Publication No. US20030032023A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: P2730P1C2
CURRENT APPLICATION NUMBER: US/09/990.711
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089512
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089538
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089600
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355

; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
 Best Local Similarity 54.0%; Pred. No. 11;
 Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCTCTGAAATGTTATGGCGCAGGCGCTGATGTCACCTGTGGAGGAGCAGGACT 1414
 DB 1998 GTCTCATTPANCARTGCAATTTGGAAGCTGTCAGTCGTCAGATTTGGAAGGCT 2057
 QY 1415 GGGATGGAAAAATAATGTGTACAGGAAACGGCAATATACATCACCAGTGGCCTGA 1474
 DB 2058 TGGATTTATCCCTGTACTTGAACATGAACACTGAATTTATGCCCTGTTTCAAGTTTGA 2117
 QY 1475 AGGAGC 1480
 DB 2118 ATGAGC 2123

RESULT 47

US-10-137-865-463
; Sequence 463, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154

; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03

; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 463
; LENGTH: 3226

; TYPE: DNA
; ORGANISM: Homo Sapien

US-10-137-865-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.0%; Pred. No. 11;

Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCTGAAAATGATTGGCGCAGGCGCTGTGACGAGCAACGCAATATACCATCAGCGGCTGGA 1414

DB 1998 GTCTCATTAAACATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057

QY 1415 GGGATGAAAAATAATGTGGTACAGCAACGCAATATACCATCAGCGGCTGGA 1474

DB 2058 TGGATTTATCCCTGTACTTGAACATGAAACTGAAATATATGCCCGTGTTCAGAGGTTGA 2117

QY 1475 AGGAGC 1480

DB 2118 ATGAGC 2123

RESULT 48

US-10-140-474-463
; Sequence 463, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162

; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 463
; LENGTH: 3226

; TYPE: DNA
; ORGANISM: Homo Sapien

US-10-140-474-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.0%; Pred. No. 11;

Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCTGAAAATGATTGGCGCAGGCGCTGTGACGAGCAACGCAATATACCATCAGCGGCTGGA 1414

DB 1998 GTCTCATTAAACATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057

QY 1415 GGGATGAAAAATAATGTGGTACAGCAACGCAATATACCATCAGCGGCTGGA 1474

DB 2058 TGGATTTATCCCTGTACTTGAACATGAAACTGAAATATATGCCCGTGTTCAGAGGTTGA 2117

QY 1475 AGGAGC 1480

DB 2118 ATGAGC 2123

RESULT 49

US-10-142-431-463

; Sequence 463, Application US/10142431

; Publication No. US20030036179A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C251

; CURRENT APPLICATION NUMBER: US/10/142,431

; CURRENT FILING DATE: 2002-05-10

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 463

; LENGTH: 3226

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-142-431-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;

Best Local Similarity 54.0%; Pred. No. 11;

Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1355 GTCCTCTGAAAATGATTGGCGCAGGCGCTGTGACGAGCAACGCAATATACCATCAGCGGCTGGA 1414

DB 1998 GTCTCATTAAACATGCAATTCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCT 2057

QY 1415 GGGTGGAAATAAATGTGTACAGGAACGGCAAAATATACCATCACCAGTGGCCTGGA 1474
|||||
Db 2058 TGGATTATCCCTCTACTTGAACATGAATTAATGCGCGTGTTCAGAGTTTGA 2117
|||||
QY 1475 AGGAGC 1480
|||||
Db 2118 ATGAGC 2123

RESULT 50.

US-10-143-114-463
; Sequence 463, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tunas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 463
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-463

Query Match 2.2%; Score 33.2; DB 9; Length 3226;
Best Local Similarity 54.0%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1355 GTCTCTGAAATAATGTTGGCGAGGCGCTGTAGTGGTGCACCTGTGGAGGAGCAGGACT 1414
|||||
Db 1998 GTCTCATTACAAATGCAATTCAGCTCGTCAGCAATGGGAAGCTGTCCATTGAAAGGCCT 2057
|||||
QY 1415 GGGTGGAAATAAATGTGTACAGGAACGGCAAAATATACCATCACCAGTGGCCTGGA 1474
|||||
Db 2058 TGGATTATCCCTGTACTTGAACATGAATTAATGCGCGTGTTCAGAGTTTGA 2117
|||||
QY 1475 AGGAGC 1480
|||||
Db 2118 ATGAGC 2123

Search completed: June 19, 2003, 06:35:34
Job time : 385.433 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:08:24 ; Search time 59.1139 Seconds
(without alignments)
7724.783 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctcgagtcggagatgaaag.....ctggaaggagcctggtcgac 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents-NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357.6	24.0	2238	2	US-08-674-887A-7
2	357.6	24.0	2238	3	US-08-951-844-7
3	357.6	24.0	2238	4	US-09-412-347-7
4	298.4	20.0	2235	1	US-08-418-782-1
5	298.4	20.0	2235	1	US-08-228-662-1
6	298.4	20.0	2235	2	US-08-852-219-1
7	298.4	20.0	2331	1	US-08-418-782-20
8	298.4	20.0	2331	1	US-08-852-219-20
9	298.4	20.0	4403765	4	US-09-103-840A-2
10	298.4	20.0	4411529	4	US-09-103-840A-1
11	274.2	18.4	2221	1	US-08-418-782-2
12	274.2	18.4	2221	1	US-08-228-662-2
13	274.2	18.4	2221	2	US-08-852-219-2
14	274.2	18.4	4794	2	US-08-459-499-8
15	274.2	18.4	4795	2	US-08-313-185-45
16	274.2	18.4	4795	3	US-09-082-614A-45
17	224.6	15.1	2262	2	US-08-674-887A-5
18	224.6	15.1	2262	3	US-08-951-844-5
19	224.6	15.1	2262	4	US-09-412-347-5
20	48	3.2	7218	1	US-08-232-463-14
21	38.8	2.6	4518	3	US-09-125-287-2
22	38.8	2.6	12839	3	US-09-125-287-1
23	33.8	2.3	579	4	US-08-991-789A-147
24	33.8	2.3	579	4	US-09-062-451-147
25	33.8	2.3	579	4	US-09-598-326-147
26	33.8	2.3	10970	3	US-08-716-351A-5
27	33.4	2.2	1074	4	US-09-134-001C-2054

28	32.4	2.2	6328	4	US-08-913-832A-1	Sequence 1, Appl
29	32.4	2.2	6328	4	US-09-249-181A-1	Sequence 1, Appl
c 30	32.4	2.2	11707	4	US-09-136-574A-1	Sequence 1, Appl
31	32.4	2.2	43795	3	US-08-742-185-101	Sequence 101, App
32	32.2	2.2	44377	2	US-08-804-227C-7	Sequence 7, Appl
33	32.2	2.2	44377	2	US-08-804-198-1	Sequence 1, Appl
c 34	32.2	2.2	2837	4	US-09-156-316-11	Sequence 11, Appl
35	31.8	2.1	2278	4	US-09-442-055-1	Sequence 1, Appl
36	31.8	2.1	3097	4	US-09-282-147-38	Sequence 38, Appl
37	31.6	2.1	2825	4	US-09-196-390-5	Sequence 5, Appl
c 38	31.6	2.1	4353	2	US-08-365-486A-18	Sequence 18, Appl
c 39	31.6	2.1	4353	4	US-08-880-342-18	Sequence 18, Appl
c 40	31.6	2.1	4780	2	US-08-365-486A-20	Sequence 20, Appl
c 41	31.6	2.1	4780	3	US-09-123-708-3	Sequence 3, Appl
c 42	31.6	2.1	4780	3	US-09-123-624-3	Sequence 3, Appl
c 43	31.6	2.1	4780	4	US-08-880-342-20	Sequence 20, Appl
c 44	31.4	2.1	3407	1	US-08-253-155A-7	Sequence 7, Appl
45	31	2.1	289	4	US-09-007-005-17	Sequence 17, Appl
46	31	2.1	289	4	US-09-244-796-17	Sequence 17, Appl
47	30.8	2.1	3153	4	US-09-175-928-9	Sequence 9, Appl
c 48	30.8	2.1	4160	4	US-09-341-678-5	Sequence 5, Appl
49	30.8	2.1	4403765	4	US-09-103-840A-2	Sequence 2, Appl
50	30.8	2.1	4411529	4	US-09-103-840A-1	Sequence 1, Appl
c 51	30.6	2.1	34053	4	US-09-453-702B-96	Sequence 96, Appl
c 52	30.6	2.1	45175	4	US-09-453-702B-116	Sequence 116, App
c 53	30.4	2.0	20199	4	US-08-961-527-6	Sequence 6, Appl
54	30.2	2.0	608	3	US-08-115-753-6	Sequence 6, Appl
c 55	30.2	2.0	849	4	US-09-134-001C-2215	Sequence 2215, App
c 56	30.2	2.0	1316	1	US-08-446-822-7	Sequence 7, Appl
c 57	30.2	2.0	1316	4	US-09-328-314-7	Sequence 7, Appl
c 58	30.2	2.0	1316	5	PCT-US93-12586-7	Sequence 7, Appl
59	30.2	2.0	5398	3	US-09-356-952-11	Sequence 11, Appl
60	30.2	2.0	9848	4	US-09-385-222A-3	Sequence 3, Appl
61	30.2	2.0	13187	4	US-09-422-936-61	Sequence 61, Appl
c 62	30.2	2.0	18627	4	US-08-961-527-113	Sequence 113, App
c 63	30	2.0	500	3	US-09-141-000-2	Sequence 2, Appl
64	29.8	2.0	729	5	PCT-US94-01149-17	Sequence 17, Appl
65	29.8	2.0	1080	4	US-09-615-192A-145	Sequence 145, App
66	29.8	2.0	1229	2	US-08-975-316-85	Sequence 85, Appl
67	29.8	2.0	1229	4	US-09-615-192A-85	Sequence 85, Appl
68	29.8	2.0	4825	6	5459251-1	Patent No. 5459251
69	29.8	2.0	5086	2	US-08-465-485A-19	Sequence 19, Appl
70	29.8	2.0	5086	2	US-08-365-486A-14	Sequence 14, Appl
71	29.8	2.0	5086	3	US-09-080-285-19	Sequence 19, Appl
72	29.8	2.0	5086	4	US-08-880-342-14	Sequence 14, Appl
73	29.8	2.0	5086	4	US-09-724-426-19	Sequence 19, Appl
74	29.8	2.0	5086	4	US-09-233-527-7	Sequence 7, Appl
75	29.8	2.0	5086	5	PCT-US93-05651-4	Sequence 4, Appl
76	29.8	2.0	5086	5	PCT-US93-06251-2	Sequence 2, Appl
77	29.8	2.0	5086	5	US-09-234-186-7	Sequence 7, Appl
78	29.8	2.0	5104	6	5506344-1	Patent No. 5506344
79	29.8	2.0	5449	4	US-09-546-990-1	Sequence 1, Appl
80	29.8	2.0	49377	1	US-08-764-233A-1	Sequence 1, Appl
81	29.6	2.0	1759	4	US-09-491-362-1	Sequence 1, Appl
82	29.6	2.0	1759	4	US-09-874-562-1	Sequence 1, Appl
83	29.6	2.0	3489	2	US-08-728-323A-1	Sequence 1, Appl
c 84	29.6	2.0	3489	4	US-09-298-568-1	Sequence 1, Appl
c 85	29.6	2.0	10348	2	US-08-457-273B-41	Sequence 41, Appl
c 86	29.6	2.0	10348	3	US-08-556-419-13	Sequence 13, Appl
c 87	29.6	2.0	10348	4	US-09-041-886-14	Sequence 14, Appl
c 88	29.6	2.0	10366	1	US-08-246-982A-5	Sequence 5, Appl
c 89	29.6	2.0	10366	1	US-08-453-265-5	Sequence 5, Appl
90	29.6	2.0	32207	2	US-08-770-379-20	Sequence 20, Appl
91	29.6	2.0	32207	4	US-08-757-669A-20	Sequence 20, Appl
92	29.6	2.0	32207	4	US-09-230-371A-20	Sequence 20, Appl
93	29.6	2.0	48908	4	US-09-453-702B-137	Sequence 137, App
c 94	29.4	2.0	409	1	US-08-400-256-37	Sequence 37, Appl
c 95	29.4	2.0	409	3	US-08-975-365-37	Sequence 37, Appl
c 96	29.4	2.0	415	1	US-08-400-256-16	Sequence 16, Appl
c 97	29.4	2.0	415	1	US-08-400-256-22	Sequence 22, Appl
c 98	29.4	2.0	415	1	US-08-400-256-25	Sequence 25, Appl
c 99	29.4	2.0	415	1	US-08-400-256-28	Sequence 28, Appl
c 100	29.4	2.0	415	1	US-08-400-256-31	Sequence 31, Appl

QY	1277	ACGTATGGCCATGGATGATGAGAGACTGTGGCCCTGTATCCGGGAGGCGATACATTTGG	13336
Db	768	CCGAATGGCCAAATGAATGACGAAGAAACCGGTGGCTCTCATATAGCGGGTGGACACACCTTCGG	827
QY	1337	TAAAGCACATGGTGCAGCGTCTCTGAAAAATGTATTGGCGCAGGCCTGTATGGTGCACC	1396
Db	828	AAAACCCATGGTGTGCGCATCGGAGAAATATGTGGGCCGAGAGCCTGCCCGCCGAGG	887
QY	1397	TGTGGAGGACGAGGACTGGGATGGAAAAATAAATGTGTACAGGAAACGGCAAAATATAC	1456
Db	888	TATTGAAGAAATGAGCCTGGGGTGGAAAAACACTACGGCACCCGGACACGGTCGGGATAC	947
QY	1457	CATCACCAGTGGCCTGGAAGAGCCTGGTGA	1488
Db	948	CATCACAGTGGACTAGAAGCGCCCTGGACCA	979
RESULT 2			
US-08-951-844-7			
; Sequence 7, Application US/08951844			
; Patent No. 6074860			
; GENERAL INFORMATION:			
; APPLICANT: Robertson et al.			
; TITLE OF INVENTION: Catalases			
; NUMBER OF SEQUENCES: 8			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,			
; ADDRESSEE: CECCHI, STEWART & OLSTEIN			
; STREET: 6 BECKER FARM ROAD			
; CITY: ROSELAND			
; STATE: NEW JERSEY			
; COUNTRY: USA			
; ZIP: 07068			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 3.5 INCH DISKETTE			
; COMPUTER: IBM PS/2			
; OPERATING SYSTEM: MS-DOS			
; SOFTWARE: WORD PERFECT 5.1			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/951,844			
; FILING DATE:			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/674,887			
; FILING DATE: July 3, 1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Charles J. Herron			
; REGISTRATION NUMBER: 28,019			
; REFERENCE/DOCKET NUMBER: 331400-55			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 201-994-1700			
; TELEFAX: 201-994-1744			
; INFORMATION FOR SEQ ID NO: 7:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2238 NUCLEOTIDES			
; TYPE: NUCLEIC ACID			
; STRANDEDNESS: SINGLE			
; TOPOLOGY: LINEAR			
; MOLECULE TYPE: DNA			
US-08-951-844-7			
Query Match 24.0%; Score 357.6; DB 3; Length 2238;			
Best Local Similarity 66.7%; Pred. No. 6.8e-107;			
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;			
QY	683	ATCAAAATCCCTGGGGGCTGATTTCGATTATGCCACCCACAGATTTCAACAGCTGGATATGA	742
Db	168	ATCGGACCCCAACGACCCGGATTTTGACTATGCCGAAGAGTTTTAAGAAGCTAGATCTGCC	227
QY	743	GGCTCTGAAAAAGATATCAAGATTTGCTGCACACTCCCGAGGATTTGTCCTCGGGA	802
Db	228	ACCGGTTAAAAAGGACTTGGCAGCGCTAATGACAGATTTACAGAGCTTGTGGCCAGCAGA	287

Db	828	AAAAACCATGGTCTGCTCCGATCGGAGAAATATGTGGCGCGAGAGCCTGCCCGCGAGG	887
QY	1397	TGTGGAGGACGAGGACTGGATGTGGAATAAATATGTGTACAGGAACGGCAAAATATAC	1456
Db	888	TATTGAAGAAATGATGACCTGGGGTGTGAAAAACACTACGGCACCGGACGGTTCGGATAC	947
QY	1457	CATCACCAGTGGCCTGGAAGGAGCCTGGTCCA	1488
Db	948	CATCACCAGTGGACTAGAAAGCGCCTGGACCA	979
RESULT 4			
US-08-418-782-1			
: Sequence 1, Application US/08418782			
: Patent No. 5658733			
: GENERAL INFORMATION:			
: APPLICANT: Cockerill, Franklin R.			
: APPLICANT: Kline, Bruce C.			
: APPLICANT: Uhl, James R.			
: TITLE OF INVENTION: Detection of Isoniazid Resistant Strains			
: TITLE OF INVENTION: of M. Tuberculosis			
: NUMBER OF SEQUENCES: 22			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Schwegman, Lundberg & Woessner			
: STREET: 3500 IDS Center			
: CITY: Minneapolis			
: STATE: MN			
: COUNTRY: USA			
: ZIP: 55402			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: Patent In Release #1.0, Version #1.25			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/418,782			
: FILING DATE:			
: CLASSIFICATION: 435			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Woessner, Warren D.			
: REGISTRATION NUMBER: 30,440			
: REFERENCE/DOCKET NUMBER: 150.141US1			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: 612-339-0331			
: TELEFAX: 612-339-3061			
: INFORMATION FOR SEQ ID NO: 1:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 2235 base pairs			
: TYPE: nucleic acid			
: STRANDEDNESS: single			
: TOPOLOGY: linear			
: MOLECULE TYPE: DNA			
US-08-418-782-1.			
Query Match 20.0%; Score 298.4; DB 1; Length 2235;			
Best Local Similarity 62.4%; Pred. No. 1.8e-87;			
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;			
QY	688	ATCCCTGGGGGCGCTGATTTCGATATGCCACACAGTTTCACACACTGGATATGAGGCTC	747
Db	176	ACCCGATGGGTGGCGGTTCGACTATCGCGGAGGTGCGGACCATCGAGCTTGACGCC	235
QY	748	TGAAAAAGATACAAAGATTTCGTGACAACTCCACAGGATTTGGTGCCTTGAACCGCTGAA	807
Db	236	TCACCGGGACATCGAGGAAGTATGACCACTCGACCGCTGGTGGCGCGGCGGACTACG	295
QY	808	GTCATTATGTCCTTTCTTTATTCGTATGGCTTGGCACGGTTCGGGAAACATACAGGACAT	867
Db	296	GCCACTAGCGGCGCTGTTTATCCGGATGGCTGGCAGCTGCGGGACCTACCGCATCC	355
QY	868	ATGATGCGCGGGAGGCCGACGTGGTCAGCAACGTTTTTGAACCGCTGAAACGCTGC	927
Db	356	ACGACGGCGCGGCGCGGGGGGCGATGACGCGGTTTCGGCGCGCTTAACAGCTGGC	415

Query Match	20.0%	Score	298.4	DB 2	Length	2235			
Best Local Similarity	62.4%	Pred.	No. 1.8e-87						
Matches	502	Conservative	0	Mismatches	296	Indels	6	Gaps	2
QY	688	ATCCCTGGGGGCGTCATTTCATTAATTCACACAGATTTCACACAGCTGGGATATGGAGGTC	747						
Db	176	ACCCGATGGGTGGCGCGTTCGACTATGCGCGGAGTTCGCACCATTCGACGTTGACGCC	235						
QY	748	TGAAAAAGATATCAAGATTTCGTGCAACACTTCCACAGATTGGTGCCCTCGGGATTATG	807						
Db	236	TGACGCGGACATCGAGGAAGTATGACCACTTCGCAGCGTGGTGGCCGCCGACTACG	295						
QY	808	GTCAATTATGGTCCTTTCTTTATTCGTATGGCTTTGSCACGGTCCGGAACATACAGGACAT	867						
Db	296	GCCACTAGCGGCCGCTGTTTATCCGGATGCGTGGCGCACGCTCGCGGCACTACCGCATCC	355						
QY	868	ATGATGGCCGGGAGGCGCCAGTGGTGTGATCAGCAAGCTTTTGAACCGCTGACACAGCTGGC	927						
Db	356	ACGACGGCCGGGGGGCGCCGGGGGGCGGATCAGCGGTTTCGCGCGCTTAACACGTGGC	415						
QY	928	CGGATAACGTTTAATCTCGATAAAGCCGCTCGATTGCTGTGGCCAGCTCAAGAAAAAATACG	987						
Db	416	CGCAACGCCAGCTTGGACAAAGCCGCGCGCTGCTGTGGCCGGTCAAGAAGAAGTACG	475						
QY	988	GCTCCAGTATTCCTGGGAGACCTGATGGTCCTGACTGGTAATGTCGCCCTTGAATCCA	1047						
Db	476	GCAAGAAGCTCTCATGGCGGACCTGATTGTTTTTCGCGGCAACTCGCGCTGGATCGA	535						
QY	1048	TGGGATTTAAACGCTGGGATTTCGTGGCGGAAGAGAAGATGACTGTGGGAGTCGGACCTGG	1107						

```

536 TGGGCTTCAAGACGCTTCGGGTTCGGGCTTCGGCGCGGGTTCGACCACTAGTGGGAGCCCGATGAGG 595
1108 TATACTGGGGGCGCTTGACAAAGACCTCTTCAGATAAACCGGGATA---AAAACGGGAAC 1164
596 TCTATTGGGCAAGGAAGCCACCTTGGCTCGCGGATGAGCGTTACAGCGGTAAAGCGGATC 655
1165 TTCAAGAAACCTTTCGGCGCCAGCAGATGGGACTTATTTATGTCAATTCCTGAAGGCCCG 1224
656 TGGAGAACCCGCTTGGCGCGGTTCAGATGGGCTGATCTACGTGAACCCGAGGGGCCGA 715
1225 GTGAAACACAGATCCTCTCGCTTCGGCTTCGCGGAAAGATATCAGGAAGCTTTTTCAGTATGG 1284
716 ACGCAACCGCGACCCCATGCGCGCGCTGCACATTCGCGAGACGTTTCGGCGCATGG 775
1285 CCATGGATGATGAGGAGACTGTGCGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
776 CCATGAACGACGTCGAACACAGCGCGCTGATCTCGCGGCTCACACTTTCGGGTAAAGCCC 835
1345 ATGGTCAGCGTCTCCTGAAATGATTTGGCCAGGGCGCTGATGGTGCACCTGTGGAGG 1404
836 ATGCGCGC---CGCCGCGCGATCTGGTGGCGCCCGAAACCGAGGCTGCTCCGCTGGAGC 892
1405 AGCAGGACTGGGATGGAATAAATATGTTGTTACAGGAACCGCAATATACCATCACCA 1464
893 AGATGGCTTGGCTTGAAGAGCTCGTATGGCACCGAACCCTGTAAGGACCGCATCACCA 952
1465 GTGGCTTGAAGAGCCTGGTCA 1488
953 GCGGCATCGAGGTCGTGAGCA 976

RESULT 7
US-08-418-782-20
; Sequence 20, Application US/08418782
; Patent No. 5658733
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,782
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.141US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2331 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:

```

NAME/KEY: CDS
LOCATION: 70...2289
US-08-418-782-20

Query Match 20.0%; Score 298.4; DB 1; Length 2331;
Best Local Similarity 62.4%; Pred. No. 1.9e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

```
QY 588 ATCCCTGGGGGGTGTATTTGATATGCCACCAAGATTTCAACAGCTGGATATGAGGCTC 747
DB 236 ACCCGATGGTGGCGGCTTCGACTATGCCGCGAGGTCGCGACCATCGACGTTGACGCC 295
QY 748 TGAATAAAGATATCAAGATTTGCTGACAACTTCCAGAGTGTGTGCTCGGATATG 807
DB 296 TGACCGGGACATCGAGAGTATGACCACTCGCAGCGTGTGTGCGCCGCCGACTACG 355
QY 808 GTCAATTATGTCCTTTCTTTATTGATGCTTGGCAGCTGCGGGAACATACAGACAT 867
DB 356 GCCACTACGGCGCGCTTTTATCCGGATGCGTGGCAGCTGCGCGCACTACCGCATCC 415
QY 868 ATGATGCGCGGGAGGCGCCAGTGTGTGAGCAACGTTTGAACCGCTGAACAGCTGGC 927
DB 416 ACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
QY 928 CGGATACGTTAATCTCGATAAAGCCCTGCTGACAACTTCCAGAGTGTGTGCTCGG 987
DB 476 CCACACCGCGAGCTTGGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 535
QY 988 GCTCCAGATATTTCTCGGGAGACGTATGATGTCCTGACTGTGTTGTCCTTGAATCCA 1047
DB 536 GCAAGAGCTCTCATGGCGGACTGATTTGCGCGCACTGCGCGCTGGAATCGA 595
QY 1048 TGGATTTAAACGCTGGGATTTGCTGGCGAGAGAGATGACTGGGAGTCGACCTGG 1107
DB 596 TGGGCTTCAAGAGCTTCGGGTTGCGGCTTCGCGCGGCTGCGACAGTGGGAGCGGATGAG 655
QY 1108 TATATGCGGCGCTGACAAAGCCCTCTTCCAGATACCGGATA---AAACGGGAAC 1164
DB 556 TCTATTGGGCAAGGAAGCCACCTTGGCTCGCGGATGAGCGTTACGCGGTAAAGCGGATC 715
QY 1165 TCCAGAAACCTCTTGGCGCCACGACAGATGGGACTTATTTATGCAATCTTGAAGCCCG 1224
DB 716 TGGAGAACCGCTGGCGCGGTCGAGATGGGCTGATCTACGTGAACCCGAGGGCGCA 775
QY 1225 GTGGAACACAGATCTCTGCTTCCGCGAAGATATCAGGGAAGCTTTTACGATGG 1284
DB 776 ACGGCAACCGGACCCCATGCGCGCGGCTGACATTCGCGAGACGTTTCGCGCATGG 835
QY 1285 CCATGGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGCGCATACATTTGTTAAAGCAC 1344
DB 836 CCATGAACGACGTCGAACAGCGCGCTGATCGCGCGGTACACTTTTCGTTAAGACC 895
QY 1345 ATGCTGACGCTCTCTTGAATAATGTTGGCGCAGGCGCTGATGTGCACTGTGGAGG 1404
DB 896 ATGGCGCG---GCCCGCGGATCTGTGTCGCGCGCAACCGAGGCTGCTCCGCTGAGC 952
QY 1405 AGCAGGACGCTGGATGAAATAATGTTGCTGACAGAAACGGCAATATACCATACCA 1464
DB 953 AGATGGGCTTGGCTGGAAGAGCTCGTATGGCAGCGGACCGGTAAGGACGCGATACCA 1012
QY 1465 GTGGCTTGAAGGAGCTGTGCTGA 1488
DB 1013 CGGCGATCGAGGTCGATGACGA 1036
```

RESULT 8

US-08-852-219-20
; Sequence 20, Application US/08852219
; Patent No. 5922575
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.

TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
TITLE OF INVENTION: of M. Tuberculosis
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.
STREET: 119 No. 5922575th Fourth Street, Ste. 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,219
FILING DATE: 07-May-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sandberg, Victoria A.
REGISTRATION NUMBER: 41,287
REFERENCE/DOCKET NUMBER: 230.00010130
TELEPHONE: 612-305-1226
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2331 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 70...2289
US-08-852-219-20

Query Match 20.0%; Score 298.4; DB 2; Length 2331;
Best Local Similarity 62.4%; Pred. No. 1.9e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

```
QY 588 ATCCCTGGGGGGTGTATTTGATATGCCACCAAGATTTCAACAGCTGGATATGAGGCTC 747
DB 236 ACCCGATGGTGGCGGCTTCGACTATGCCGCGAGGTCGCGACCATCGACGTTGACGCC 295
QY 748 TGAATAAAGATATCAAGATTTGCTGACAACTTCCAGAGTGTGTGCTCGGATATG 807
DB 296 TGACCGGGACATCGAGAGTATGACCACTCGCAGCGTGTGTGCGCCGCCGACTACG 355
QY 808 GTCAATTATGTCCTTTCTTTATTGATGCTTGGCAGCTGCGGGAACATACAGACAT 867
DB 356 GCCACTACGGCGCGCTTTTATCCGGATGCGTGGCAGCTGCGCGCACTACCGCATCC 415
QY 868 ATGATGCGCGGGAGGCGCCAGTGTGTGAGCAACGTTTGAACCGCTGAACAGCTGGC 927
DB 416 ACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
QY 928 CGGATACGTTAATCTCGATAAAGCCCTGCTGACAACTTCCAGAGTGTGTGCTCGG 987
DB 476 CCACACCGCGAGCTTGGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 535
QY 988 GCTCCAGATATTTCTCGGGAGACGTATGATGTCCTGACTGTGTTGTCCTTGAATCCA 1047
DB 536 GCAAGAGCTCTCATGGCGGACTGATTTGCGCGCACTGCGCGCTGGAATCGA 595
QY 1048 TGGATTTAAACGCTGGGATTTGCTGGCGAGAGAGATGACTGGGAGTCGACCTGG 1107
DB 596 TGGGCTTCAAGAGCTTCGGGTTGCGGCTTCGCGCGGCTGCGACAGTGGGAGCGGATGAG 655
QY 1108 TATATGCGGCGCTGACAAAGCCCTCTTCCAGATACCGGATA---AAACGGGAAC 1164
DB 556 TCTATTGGGCAAGGAAGCCACCTTGGCTCGCGGATGAGCGTTACGCGGTAAAGCGGATC 715
```



```

QY 1165 TTCAGAACCTCTTGGCGCACGACAGATGGGACTTATTTATGTCATCTCTGAAGGCCCG 1224
Db 716 TGGAGAACCCGCTGGCGCGGTGCAGATGGGTGATCTACGTGAACCCGGAGGGCCGA 775
QY 1225 GTGGAACACAGATCTCTGGCTTCGGGAGATATCAGGAGCTTTTTCACGTATGG 1284
Db 776 ACGCAACCCGGACCCCATGGCGCGGTGCACATTCGCGAGACGTTTCGGCGCATGG 835
QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGATACATTTGCTAAAGCAC 1344
Db 836 CCATGAGACGCTCGAAACAGCGGCTGATGTGCGCGGTGCACATTTTCGCTAAGACCC 895
QY 1345 ATGGTCAGGCTCTCTGAAATGATTTGGCGCAGGCGCTGATGTGCACCTGTGGAGG 1404
Db 896 ATGGCGCGG---CCCGCGCGGATCTGTGCGCGCGGCAACCGAGGCTGCTCGCTGGAGC 952
QY 1405 AGCAGGAGACTGGGATGGAATAATGTTGGTACAGGAACGGCAATATACCATCACC 1464
Db 953 AGATGGCTTGGCTGGAAGAGCTGTATGGCACCAGGAAACCGGTAAGGACGCGATCACC 1012
QY 1465 GTGGCTGGAAGAGCCTCGTCTGA 1488
Db 1013 GCGGCATCGAGTCTGTATGGACGA 1036

RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 20.0%; Score 298.4; DB 4; Length 4403765;
Best Local Similarity 62.4%; Pred. No. 1.8e-85;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGCGCTGATTTGATTCATCCACAGATTTCAACAGCTGGATGAGGCTC 747
Db 2153203 ACCCGATGGTGGCGGTTCGACTATGCCGGGAGGTGCGACCATCGACTTCAGCGCC 2153144
QY 748 TGAAGAAGATATCAAGATTTGTCGACACTTCCAGGATTTGGTCCCTCGCGGATATG 807
Db 2153143 TGACCGGGACATCAGGAAGTGTATGACCACTTCGACGCGGTGTGGCGCGGACTACG 2153084
QY 808 GTCATATGTCCTTTCTTTATTCGTATGCTTGGACGCTGCGGAGATATGATGAGACAT 867
Db 2153083 GCCACTAGGCGGCTGTTATTCGGATGGGTGGACGCTGCGGCACTACCGCATCC 2153024
QY 868 ATGATGCGCGGGAGCGCCAGTGTGGTTCAGCAACCTTTTGAACCGCTGAACAGCTGCG 927
Db 2153023 ACGACGCGCGCGCGCGCGGCGGATGCGAGGTTTCGCGCGCTTAACAGCTGCG 2152964
QY 928 CGGATACGTTATCTGGATAAAGCCGCTGATGCTGTGGCCAGTCAAGAAAAAATACG 987

```

```

Db 2152963 CGCAACAGCCAGCTTGGACAGGCGCGCGCTGTGCGCGGTCAAGAAGATACG 2152904
QY 988 GCTCCAGTATTTCCTGGGAGACCTGATGGTCTGACTGGTAATGTTGCCCTTCAATCCA 1047
Db 2152903 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGGCACTGCGCGCTGGAATCGA 2152844
QY 1048 TGGGATTTAAACGCTGGGATTTCTGCGGGAAGAAGATGATCTGGAGTTCGAGCTGG 1107
Db 2152843 TGGGCTTCAAGACCTTCGGGTTCGGCTTCGCGCGGTTCGACCACTGCGAGCGCGATGAGG 2152784
QY 1108 TATACTGGGGCCCTGACAACAAGCCTCTTCAGATAACCGGATA---AAAACGGGAAC 1164
Db 2152783 TCTATTGGGCAAGAAGACCACTTGGCTCGCGGATGAGCGTTACAGCGGTAAGCGGATC 2152724
QY 1165 TTCAGAAACCTTTTTCGCGCACGATGGGACTTATTTATGTCATCTTGAAGGCCCG 1224
Db 2152723 TGGAGAACCCGCTGCGCGGTGCAGATGGGCTGATCTACGTGACCCCGAGGGGCCGA 2152664
QY 1225 GTGAAACACAGATCTCTGCTTCCGCGAAGATATCAGGAGAGCTTTTTCAGTATGG 1284
Db 2152663 ACGCAACCCGCGACCCCATGGCGCGGTGCACATTCGCGAGACGTTTCGCGCGATGG 2152604
QY 1285 CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
Db 2152603 CCATGAACGAGCTCGAACAACAGCGGCTGATCTCGCGGTGCACATTTTCGGAACCC 2152544
QY 1345 ATGGTCAGCCTCTCTGAAATGATTTGGCGCAGGCGCTGATGTGACCTGTGGAGG 1404
Db 2152543 ATGGCGCG---GCCGCGCGATCTGGTGGCGCCCGAAACCCGAGGCTGCTCCGCTGAGC 2152487
QY 1405 AGCAGGAGCTGGGATGGAATAATAATGTGTACAGGAACGCAATATACCATCACC 1464
Db 2152486 AGATGGCTTGGCTGGAAGAGCTCGTATGGCACCGGACCGGTAAGGACGCGATCACC 2152427
QY 1465 GTGGCTGGAAGAGCCTGTGCTGA 1488
Db 2152426 GCGGCATCGAGTCTGTATGGACGA 2152403

RESULT 10
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 20.0%; Score 298.4; DB 4; Length 4411529;
Best Local Similarity 62.4%; Pred. No. 1.8e-85;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGCGCTGATTTGATTCATCCACAGATTTCAACAGCTGGATGAGGCTC 747
Db 2155943 ACCCGATGGTGGCGGTTCGACTATGCGCGGAGGTGCGGACCATCGACTTACGCGCC 2155884
QY 748 TGAAAAAAATATCAAGATTTGTCGACAACTTCCAGGATTTGGTCCCTCGGATATG 807

```

Db 2155883 TGACGGGACATCGAGGAAGTGTGACCACTCGCAGCGTGGTGGCCCGCGACTACG 2155824
QY 808 GTCAATATGTCCTTTCTTATTCGTATGCTTGGACAGGTGCGGAAACATACAGACAT 867
Db 2155823 GCCACTACGGCGCGCTTTATCCGATGGCTGGCAGCTGCGGCACCTACCGCATCC 2155764
QY 868 ATGATGCGCGGGAGCGCGCAGTGGTGTGCTGAGCAAGCTTTTGAACCGCTGAACGCTGGC 927
Db 2155763 ACACG 2155704
QY 928 CGGATACGCTTAATCTGGATAAAGCCCGCTGATGTGTGCGCCAGTCAAGAAAAATACG 987
Db 2155703 CCACAAAGCCAGCTTGGCAAGCGCGCGCGCGCTGTGTGCGCGCTCAAGAAAGTACG 2155644
QY 988 GCTCCAGTATTTCTGCGGAGACCTGATGTGCTGACTGTATGTGCTTGGCTTGAATCCA 1047
Db 2155643 GCAGAAGCTCTCATGGCGGACCTGATTTGCTGCGGCAACTGCGCGCTGGAATCGA 2155584
QY 1048 TGGGATTTAAACGCTGGGATTTGCTGCGGGAAGAGATGACTGGGAGTGGACCTGG 1107
Db 2155583 TGGCTTCAAGACGTTTCGGGTTTCGGCTTCCGCGCGGTGCGACAGTGGGAGCCCGATGAGG 2155524
QY 1108 TATATGCGGCGCTGCAACAAGCCTCTTTCAGATAACCGGGATA---AAAACGGGAAC 1164
Db 2155523 TCTATTGGCGCAAGGAGCCACCTGGCTCGCGCATGAGCGTTACAGCGGTAAAGCGATC 2155464
QY 1165 TTCAAGAACCTCTTGGCGGCAGCAGAGTGGAGCTTATTTATGTCAATCTTGAAGGCCCG 1224
Db 2155463 TGGAGAACCCGCTGGCGCGCGCTGACATGGGCTGATCTACGTGAACCCCGGAGGCGCA 2155404
QY 1225 GTGAAACACAGATCTCTGCTGCTCCGCGAAGATATCAGGAGCTTTTTCAGTATGG 1284
Db 2155403 ACGGCAACCGGAGCCCATGGCGCGCGCTGACATTCGCGAGAGCTTTTCGGCGCATGG 2155344
QY 1285 CCATGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
Db 2155343 CCATGAACGAGCTGCAACAGCGCGCTGATCGTGGCGGTACACACTTTCGGTAAAGACC 2155284
QY 1345 ATGGTGCAGGCTCTCTGAAAATGATTTGGCGAGGCGCTGATGTGCACTGTGGAGG 1404
Db 2155283 ATGGCGCG---GCCCGCGCGATCTGGTTCGCGCCCGAACCCGAGGCTGCTCCGCTGGAGC 2155227
QY 1405 AGCAGGAGCTGGGATGAAATAAATGTGCTGAGGAAACGCAATATACCATCACCA 1464
Db 2155226 AGATGGCTTGGCTGGAAGAGCTGCTGATGGCACCAGGACCGGTAAGGACCGCATCACA 2155167
QY 1465 GTGGCTTGAAGGAGCTGTGCTGA 1488
Db 2155166 GCGGCATCGAGGTCGTATGCACGA 2155143

RESULT 11

US-08-418-782-2

; Sequence 2, Application US/08418782

; Patent No. 5658733

; GENERAL INFORMATION:

; APPLICANT: Cockerill, Franklin R.

; APPLICANT: Kline, Bruce C.

; APPLICANT: Uhl, James R.

; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESS: Schwegman, Lundberg & Woessner

; STREET: 3500 IDS Center

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/418,782

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Woessner, Warren D.

; REGISTRATION NUMBER: 30,440

; REFERENCE/DOCKET NUMBER: 150.141US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-339-0331

; TELEFAX: 612-339-0361

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2221 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

US-08-418-782-2

Query Match 18.4%; Score 274.2; DB 1; Length 2221;

Best Local Similarity 61.4%; Pred. No. 1.6e-79;

Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;

QY 688 ATCCCTGGGGGCTGATTTTGTATATGCTGACCAAGATTTTCAACAGCTGGGATATGAGGCTC 747

Db 176 ACCCGATGGGTGGGGCTTCGACTATGCCGCGAGGTGCGGACAGTCGACTTTCGACGCC 235

QY 748 TGAATAAAGATATCAAGATTTCTGACAACTTCCAGGATTTGTCGCCCTGCGGATATG 807

Db 236 TGACGGGAGATCGAGGAAGTGTGACCACTTCGCAAGCTGCGGCGCTGGCGCGGACTACG 295

QY 808 GTCAATATGCTCTTCTTTATTCGTATGCTTGGCAGCTGCGGAAACATACAGGACAT 867

Db 296 GCCACTACGGCGCGCTTTATCCGATGCGTGGCAGCTGCGGACCTACCGCATCC 355

QY 868 ATGATGGCGGGAGGCGCGAGTGTGTGTCAGCAAGCTTTTGAACCGCTGAAAGCTGGC 927

Db 356 ACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCGCGCTTAAACAGCTGGC 415

QY 928 CGGATACGCTTAATCTGTGATAAAGCCGCTGATTTGCTGCGCAGTCAAGAAAAATACG 987

Db 416 CCACAAAGCAGCTTGGACAGCGCGCGCGCTGCTGTCGCGGCTCAGAAAGTACG 475

QY 988 GTCCAGTATTTCTTGGGAGACCTGATGTCTCTGACTGTAA--TGTTCCTTGAATC 1045

Db 476 GCAAGAAGCTCTCATGGCGGACCTGATTTTTCGCGCGCAACCGCTGCGCTCGGAATC 535

QY 1046 CATGGGATTTAAACGCTGGGATTTGCTGCGGAGAGAGATGACTGGGAGTGGACCT 1105

Db 536 GATGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGG--GCGTCGACAGTGGGAGACCGATGA 593

QY 1106 GGTATCTGCGGCGCTGACAAAGCCCTTTCGAGATAACCGGATAAAGGAAACCT 1165

Db 594 GGTCTATTTGGGCAAGAACCCACCTGCGCTGCGGATGCGGTTTACAGCTTAAGCTATCT 653

QY 1166 TCAGAAACCTCTTTCGCGCGCAGGAGTGGGACTTATTTATGCTCAATCTCTGAAGGCCCGG 1225

Db 654 GGAGAACCCGCTGCGCGCGTGCAGATGGGCTGATCTAGCTGACCCGAGGCGCGGAA 713

QY 1226 TGAAGAACAGATCTCTGCTGCTTCCGCGAAAGATATACAGGAAGCTTTTTCAGTATGCG 1285

Db 714 CGGCAACCCGCGGACCCATGCGCGCGGTTCGACATTCGCGAGACGTTTTCGCGCATGCG 773

QY 1286 CATGGATGATGAGGAGACTTGGCCCTGATCGCGGAGGCGCATACATTTGTTAAACACA 1345

Db 774 CATGAAGGAGCTCGAAACAGCGGCGCTGATCTGCGCGGCTCACACTTTCGCTAAGACCA 833

QY 1346 TGGTGCAGCGCTCTCTGAAATGTTATTTGGCGAGGCGCTGATGTTGCTGACCTCTGGAGGA 1405

Db 834 TGGCGCG---GCCCGCGCATCTGGTTCGCGCGCGCGCGCGCGCGCTCCGCTGGAGCA 890

QY 1406 GCAGGACTGGATGGAATAAATGTTGTACAGGAACGGCAAAATATACCATCACAG 1465
 Db 891 GATGGCTTGGCTGGAGAGCTCGTATGGCACCGGAACGGTAAGGACGGATCACAG 950
 QY 1466 TGGCTGCAAGGAGCCTGGTCGA 1488
 Db 951 CGGATCGAGTGTATGGACGA 973

RESULT 12
 US-08-228-662-2
 ; Sequence 2, Application US/08228662
 ; Patent No. 5688639
 ; GENERAL INFORMATION:
 ; APPLICANT: COCKERILL, FRANKLIN R.
 ; APPLICANT: KLINE, BRUCE C.
 ; APPLICANT: UHL, JAMES R.
 ; TITLE OF INVENTION: DETECTION OF ISONIAZID RESISTANT STRAINS
 ; TITLE OF INVENTION: OF M. TUBERCULOSIS
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCHWEGMAN, LUNDBERG & WOESSNER, P. A.
 ; STREET: 3500 IDS CENTER
 ; CITY: MINNEAPOLIS
 ; STATE: MINNESOTA
 ; COUNTRY: USA
 ; ZIP: 55402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/228,662
 ; FILING DATE: 18-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WOESSNER, WARREN D.
 ; REGISTRATION NUMBER: 30,440
 ; REFERENCE/DOCKET NUMBER: 150.123US01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-339-0331
 ; TELEFAX: 612-339-3061
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 221 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-228-662-2

Query Match 18.4%; Score 274.2; DB 1; Length 2221;
 Best Local Similarity 61.4%; Pred. No. 1.6e-79;
 Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;

QY 688 ATCCCTGGGGCGCTGATTTGATTATGCCACAGATTTCACACCTGGATGGAGCTC 747
 Db 176 ACCCGATGGTGGCGCTGACTATGCCGGGAGTGGACAGTCGACTTGGACGCC 235
 QY 748 TGAAGAAGATATCAAGATTGCTGACAACTCCAGGATTGTCCTCGCGGATTATG 807
 Db 236 TCACCGGGACATCAGGAAGTATGACCACTCGCAGCGCTGTGGCGCGCGACTACG 295
 QY 808 GTCAATTATGTCCTTCTTTATTCGATGGCTGGCAGCTGGCGGACATACAGGACAT 867
 Db 296 GCCACTAGGCCCGCTGTTATCCGGATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 355
 QY 868 ATGATGGCGGGAGCGCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 927
 Db 356 ACGACGGCGGGCG 415
 QY 928 CGGATACGTTATCTGGATAAGCCCGTGCATTTGCTGTGGCCAGTCAAGAAAAATACG 987

Db 416 CGACACGCCAGCTTGGCAAGCGCGCGCTGCTGTGGCCGCTCAAGAAGAACTACG 475
 QY 988 GCTCCAGTATTCCTGGGGAGACCTGATGTCCTGACTGGTAA--TCTTCCCTTTGAATC 1045
 Db 476 GCAAGAAGCTCTCATGGCGGACCTGATTTTCGCGCGCAACCGCTGCTGGTGGAACT 535
 QY 1046 CATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAAAGATGACTGGGAGTCGGACCT 1105
 Db 536 GATGGGCTTCAAGAGCTTGGGTTCCGCTTCGG--CGTCCAGCAGTGGGAGACCGATGA 593
 QY 1106 GGTATCTGGGGGCTGACACAAAGCCTCTTGCAGATAACCGGATATAAACCGGAACCT 1165
 Db 594 GGTCTATTGGGCAAGAACCCCTCGCTCGCGGATGACGGTTACAGCCTAAGCGATCT 653
 QY 1166 TCAGAAACCTCTTCCGCCACGACAGATGGGACTTATTATCTCAATCTGAAGGCCCGG 1225
 Db 654 GGAGAACCCGCTGGCGGCTGCAGATGGGCTGATCTAGTGAACCGGAGGCGCGAA 713
 QY 1226 TGGAAACACAGATCTCTGCTTCCGCGGAAGATATCAGGGAAGCTTTTTCAGTATGGC 1285
 Db 714 CGCAACCCGACCCCATGGCGCGGCTCGACATTCGCGAGACGTTTCGGCGCATGGC 773
 QY 1286 CATGATGATGAGGAGACTGTGCGCTGTATCGCGGAGGCGCATATTTGTTAAACACA 1345
 Db 774 CATGAACGACGTGAAACAGCGCGCTGATCGTGGCGGTACACACTTTCGGTAAGACCA 833
 QY 1346 TGGTGCAGCTCTCTGTAATAATGATTTGGCGAGGCGCTGATGGTGCACCTGTGGAGGA 1405
 Db 834 TGGCGCG---GCCCGCGCATCTGTCGCGCCCGCAACCGAGGCTGCTCCGCTGGAGCA 890
 QY 1406 GCAGGGACTGGGATGGAATAAATAATGCTGTCAGGAACGGCAAAATATACCATCACAG 1465
 Db 891 GATGGGCTTGGCTGGAGAGCTCGTATGCCACCGAACCCTGGTAAGACCGATCACCA 950
 QY 1466 TGGCTTGAAGAGCCTGTGTCGA 1488
 Db 951 CGGATCGAGTGTATGGACGA 973

RESULT 13
 US-08-852-219-2
 ; Sequence 2, Application US/08852219
 ; Patent No. 5922575
 ; GENERAL INFORMATION:
 ; APPLICANT: COCKERILL, Franklin R.
 ; APPLICANT: Kline, Bruce C.
 ; APPLICANT: Uhl, James R.
 ; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
 ; TITLE OF INVENTION: of M. Tuberculosis
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.
 ; STREET: 119 No. 5922575th Fourth Street, Ste. 203
 ; CITY: Minneapolis
 ; STATE: MN
 ; COUNTRY: USA
 ; ZIP: 55401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/852,219
 ; FILING DATE: 07-May-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sandberg, Victoria A.
 ; REGISTRATION NUMBER: 41,287
 ; REFERENCE/DOCKET NUMBER: 230.00010130
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-305-1226


```
Db      2920 CGGCATCGAGGTCTGTATGGACGA 2942
|||||
RESULT 16
US-09-082-614A-45
; Sequence 45, Application US/09082614A
; Patent No. 6124098
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,614A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/313,185
; FILING DATE: 12-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-082-614A-45
Query Match 18.4%; Score 274.2; DB 3; Length 4795;
Best Local Similarity 61.4%; Pred. No. 2.5e-79;
Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;
QY      688 ATCCCTGGGGGCTGATTTTGGATATGCGCACCAAGATTTCACAGCTGGATATGGAGCTC 747
Db      2145 ACCCGATGGGTGGCGGCTTCGACTATGCGCGGAGGTCGCGACCACTTGGAGCCC 2204
QY      748 TGAAGAAATATCAAGATTGCTGACAACTTCCACAGATTGGTGCCTCGGATTATG 807
Db      2205 TGACGGGACATCGAGGAAGTATGACCACTCGACGCTGGTGGCCCGCGACTACG 2264
QY      808 GTCAATATGTCCTTTCTTTATGATGCTTGGCAGCTGGCGGAACATACAGGACAT 867
Db      2265 GCCACTACGGCGCTGTTTATCGGATGCGTGGCAGCTGCGGCACTACCGCATCC 2324
QY      868 ATGATGCCGGGAGGCGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 927
|||||
Db      2325 ACGACGGCCCGCGGGCGCGGGCGGCGGATGACGAGGGTTCCGCGCGCTTAAACAGCTGGC 2384
QY      928 CGGATAACGTTAATCTGGATAAAGCCCGTCGATTGCTGTGTGGCCAGTCAAGAAAAAATACG 987
Db      2385 CCACAAAGCCAGCTTGGACAAGCGCGCCGCTGCTGTGGCCGTTCAAGAAGATACG 2444
QY      988 GCTCCAGTATTCTCTGGGAGAGCTGATGTCTCTGACTGTGTA--TGTTCCTCTTGAATC 1045
Db      2445 GCAAGAAGCTCTCATGGCGGACCTGATTCTTTTCGCGCGCAACCCGCTGCGCTCGGAATC 2504
QY      1046 CATGGGATTTAAACAGCTGGATTCTGTCGCGAAGAGATGACTGGGAGTCGGACCT 1105
Db      2505 GATGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGG--GGTCGACAGTGGGAGACCGGATGA 2562
QY      1106 GTTATCTGGGGCCCTGACAAACAGCCTCTTTCAGATAAACCGGATAAAAAACGGGAACCT 1165
Db      2563 GGTCTATTGGGCGCAAGCAAGCCACCTGGCTCGCGATGACGGTTACAGCGTAAGCGATCT 2622
QY      1166 TCAGAAACCTCTTTCGCGCCAGCAGATGGACCTTATTATGTCAATCTCAAGCCCGCG 1225
Db      2623 GGAGAACCCGCTGCGCGCGGTGCAGATGGGGCTGATCTAGCTGAACCCGAGCGCGCAA 2682
QY      1226 TGAAGAACCAAGATCCTCTGCTTCCGCGAAGATATCAGGAAAGCTTTTTCAGTATGGC 1285
Db      2683 CGCAACCCGACCCCATGCGCGGCGGTTCGACATTCGCGAGACGTTTCGGCGCATGGC 2742
QY      1286 CATGGATGATGAGGAGACTGTGCGCCCTGATCGCGGAGGAGCATACATTTGGTAAAGACA 1345
Db      2743 CATGAACGAGCTCGAAACAGCGCGCTGATCGTGGCGGTTCACACTTTTCGGTAAGACCCA 2802
QY      1346 TGGTGACGCTCTCTTGAAATAATGATTGGCGGAGGCGCTGTTGGTGCACCTCTGGAGGA 1405
Db      2803 TGGCGCGG---GCCCGCGCGATCTGCTGCGCCCGCAACCCGAGGCTGCTCCGCTGGAGCA 2859
QY      1406 GCAGGGACTGGGATGGAATAATAATGCTGTACAGGAAACCGCAATATATACCATCACAG 1465
Db      2860 GATGGGCTTGGCTGGAGAGCTGCTGATGACCGGAACCGGTAAAGACCGGATCACCAG 2919
QY      1466 TGGCCTGGAAGAGACCTTGGTCGA 1488
Db      2920 CGGCATCGAGGTCTGTATGGACGA 2942
RESULT 17
US-08-674-887A-5
; Sequence 5, Application US/08674887A
; Patent No. 5939300
; GENERAL INFORMATION:
; APPLICANT: Robertson, Dan E.
; APPLICANT: Sanyal, Indrajit
; APPLICANT: Adhikari, Robert S.
; TITLE OF INVENTION: CATALASES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,887A
; FILING DATE: 03-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Ph.D., Lisa A.
```

REGISTRATION NUMBER: 38,347	15.1%;	Score 224.6;	DB 2;	Length 2262;
REFERENCE/DOCKET NUMBER: 09015/002001	Best Local Similarity	57.4%;	Pred. 3.1e-63;	
TELECOMMUNICATION INFORMATION:	Matches 476;	Conservative 0;	Mismatches 329;	Indels 24; Gaps 3;
TELEPHONE: 619/678-5070				
TELEFAX: 619/678-5099				
INFORMATION FOR SEQ ID NO: 5:				
SEQUENCE CHARACTERISTICS:				
LENGTH: 2262 base pairs				
TYPE: nucleic acid				
STRANDEDNESS: double				
TOPOLOGY: linear				
MOLECULE TYPE: cdna				
FEATURE:				
NAME/KEY: Coding Sequence				
LOCATION: 1...2259				
US-08-674-887A-5				

QY	682	AATCAAAATCCCTGGGGCGCTGATTTTGATATGCACACAGATTTCAACAGCTGATGATGG	741
DB	218	AATCAGACCCGATGGATCGGATTTCAACTACCTGGAAGAGTACGCAAGCTCGATTTCG	277
QY	742	AGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCCAGGATGTGGCCCTCGG	801
DB	278	ACGCGCTCGAAGAAAGATGTCACCGCTTGATACCGATAGCCAAGAGTGGTGGCCGCTG	337
QY	802	ATTATGTCATTATGTGCTTCTTTTATTGCTATGGCTTGCGAGGTGCGGACATACA	861
DB	338	ACTGGGGCACTACGGCGGTTTGATGATCGTATGGCTTGGCACTCCGCTGGCACCTACC	397
QY	862	GGACATATGATGGCCGGGAGGCCAGTGGTGTGCTAGCAACGCTTTTGAACCGCTGAACA	921
DB	398	GTATTGTGATGGCGCTGGGGCGGTGGTACCAGCAAGCCAGCGCTTTGCACGCTCACT	457
QY	922	GCTGGCCGGATTAAGTTAATCTGGATAAAGCCGCTGATTTGCTGTGGCCAGTCAAGAAA	981
DB	458	CTTGGCCGGAACGCTCAGCGCTGGATAAAGCGCGCGCTGCTGTGGCCGATCAAGAAGA	517
QY	982	AATACGGCTCCAGTATTTCCTGGGGAGACCTGATGTCTGTGCTTAATGTTGCCCTTG	1041
DB	518	AGTACGGCAACAAATCAGCTGGCAGACCTGATTTCTGGCTGGCACCGTGGCTTATG	577
QY	1042	AATCCATGGATTAAACCGCTGGGATTCTCTGGCGGAAGAGAAGATGACTGGGAGTCGG	1101
DB	578	AGTCCATGGGCTTACCTGTCTAGCGGCTCTCTTTTCGGCCGCTCGATATTTGGGAACCG	637
QY	1102	ACCTGG---TATACTGGGGCCGTGACAAACAGCCTCTTGCAAT-----	1142
DB	638	AAAAGATATCTACTGGGTGACAAAAGAGTGGCTGGCACCTCTCTGACGAACGCTACG	697
QY	1143	--AACCGGGATAAAAACGGGAACCTTTCAGAAACCTTTGGCCGACGACAGATGGGACTTA	1200
DB	698	CGCAGCTGAAACAGCCAGAGACCATGAAAACCGCTGGCGGTGTCTCAAATGGGTCTGA	757
QY	1201	TTTATCTCAATCTGAAGCCCCGGTGGAAAACCAAGATCTCTGTGGCTTCGGGAAAGATA	1260
DB	758	TCATATGTGAACCCGGAAGTGTGTAACGGCCACCTGATCGCTGAGAAGCCGACAGCAG	817
QY	1261	TCAGGGAAGCTTTTTCACGTATGGCCATGGATGATCAGGAGACTGTGGCCCTGATCGCG	1320
DB	818	TACTTTGAACCTTCGCCCGTATGGCGATGACAGCAAGAAAACCGCAGCCCTCACAGCTG	877
QY	1321	GAGGCGATACATTGTTGTAAGCACATPGGTGCAGCGTCTCTTGAAAATATGTTGGCGCAG	1380
DB	878	CGCGCCACACCGCTCGGTAATTTGTACGGTAATTTGGCAATGCCCTCTGCGT---TAGCCCTG	934
QY	1381	GGCTCATGTGCACCTGTGGAGGACACGGGACTGGGATGGAAAATAAATATGTTGTACAG	1440
DB	935	ACCCAAAGCCCTGACGTTGAAACCCAGGGCTTAGTTTGGGGCAACCCCAACATGCGAG	994

982	QY	AAATACGGCTCCAGTATTCTCTGGGAGACCTGATGTCCTCTGACTGTGTAATGTTGCCCTTG	1041
518	Db	AGTACGGCAACAAATACAGCTGGCAGACCTGATGATCTTGCTGGTCGACCGTGGCTTATG	577
1042	QY	AATCCATGGATTAAACCGCTGGGATTTCCTGGCCGAAGAGAAGATGACTGGGAGTCGG	1101
578	Db	AGTCCATGGCTTTACCTGCTTACGGCTTCTCTTTCGGCCGCGTCGATATTTGGGAACCCG	637
1102	QY	ACCTGG---TATACTTGGGGCCTTGACAAAGAGCTCTTGCGAT-----	1142
638	Db	AAAAAGATATCTACTGGGTGACAAAGAGTGGCTGGCACCTTCTGACGAACGCTACG	697
1143	QY	--AACCGGATAAAAACGGGAACTTCAGAAAACCTCTGCCGCCACGACAGATGGAGCTTA	1200
698	Db	GCAGCTGAACAAGCCAGACAGACCATGAAAACCCGCTGGCGCTGTCCAAATGGGTCTGA	757
1201	QY	TTTATGTCAAATCCTGAAGGCCCGGTGGAAAACCAAGATCCTCTGGCTTCGCGGAAAGATA	1260
758	Db	TCATATGAACCCGGGAAGTGTTTAACGGCCACCTGATCCGCTTGAGAACCGCACAGCAGG	817
1261	QY	TCAGGGAAGCTTTTTCACGTATGGCCATGATGAGGAGACTGTGGCCCTCATCGCGG	1320
818	Db	TACTTTAAACCTTCGCCCCGTATGGCCATGAACGACGAAACCCGACGCCCTCACAGCTG	877
1321	QY	GAGGGCATACATTTGGTAAAGACATAGGTGCACGCTCTCCTGAAAAATGTATTGGCGCAG	1380
878	Db	GGGGCCACACCGTCGTTATTTGCACGGTAATGGCAATGCCTCTGGGT---TAGCCCTTG	934
1381	QY	GGCCTGATGGTCACCTGTGGAGGACGAGGACTGGGATGGAAAAATAAATGTGGTACAG	1440
935	Db	ACCCAAAAGCCTCTGACGTTGAAAACACAGGCCTTAGGTTGGGCAACCCCAACATGCAGG	994
1441	QY	GAACGGCAATAATACCATCACAGTGGCCTGGGAAGAGCCTGGTGCAG	1489
995	Db	GCAAGCAAGCAACGGCGTACCTCGGGTATCGAAGTGTCTGGAGCAC	1043

RESULT 19

US-09-412-347-5
Sequence 5, Application US/09412347
Patent No. 6410290
GENERAL INFORMATION:
APPLICANT: Robertson, Dan E.
APPLICANT: Sanyal, Indrajit
APPLICANT: Adhikari, Robert S.
TITLE OF INVENTION: CATALASES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/412,347
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/674,887
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09015/0020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:			
LENGTH:	2262 base pairs		
TYPE:	nucleic acid		
STRANDEDNESS:	double		
TOPOLOGY:	linear		
MOLECULE TYPE:	cdna		
FEATURE:			
NAME/KEY:	Coding Sequence		
LOCATION:	1...2259		
US-09-412-347-5			
Query Match	15.1%;	Score 224.6;	DB 4; Length 2262;
Best Local Similarity	57.4%;	Pred. No. 3.1e-63;	
Matches 476;	Conservative 0;	Mismatches 329;	Indels 24; Gaps 3;
QY	682	AATCAATCCCTGGGGGCTGATTTTGATTATGCCACCAGATTTCAACAGCTGATGATGG	741
DB	218	ATACAGACCCGATGGATTCGGATTTCACTACCGTGAAGNAGTACCGACGCTCGATTTTCG	277
QY	742	AGGCTCTGAAAAAAGATATCAAGATTTGCTGACAACTTCCCAAGATTTGGTGCCTTCGCG	801
DB	278	ACGCGCTGAAGAAGATGTCACGCGCTTGATGACCGATAGCCAAGAGTGGTGGCCGCTG	337
QY	802	ATTATGGTCATTATGTCCTTCTTTATTCGTATGCTTGGCAAGTGGCGGAACATACA	861
DB	338	ACTGGGGGCACTACGCGCGTTTGATGATCCGTATGCTTTGGCACTCCGCTGGCACCTACC	397
QY	862	GGACATATGATGCCGGGAGGCGCAGTGGTGTCAGCAACGTTTTTGAACCGCTGAACA	921
DB	398	GTAATTGCTGATGGCCGTGGGGCGGTGGTACCGGAAGCAGCGCTTTGACCGCTCAACT	457
QY	922	GCTGGCCGGATACGCTTAATCTGGATAAAGCCCGTCGATTGCTGTGCCAGTCAAGAAA	981
DB	458	CCTGGCCGACACGCTCAGCTGGATGAAGCGCGCTCTGCTGTGGCGATCAAGAAGA	517
QY	982	AATACGGCTCCAGTATTTCTTGGGGAGACCTGATGTCGACTGTTAATGTTGCCCTTG	1041
DB	518	AGTACGGCAACAATAATCAGCTGGGCAGACCTGATGATTTCTGGCTGGCACCGTGGCTTATG	577
QY	1042	AATCCATGGGATTTAAACCGCTGGGATTTCTGGCGGAAGAGAAGATGACTGGGAGTCGG	1101
DB	578	AGTCCATGGGCTTACCTGCTTACGGCTTCTTTTCGGCCGCTCGATATTTGGGAACCG	637
QY	1102	ACCTGG---TATACTGGGGCCCTGACAAAGCCTCTTGAGAT-----	1142
DB	638	AAAAAGATATCTACTGGGTGACAAAAAGAGTGGCTGGCACCTCTGACGAACGCTACG	697
QY	1143	--AACCGGATAAAAACGGGAACCTTCAGAAAACCTCTGCCGCCACGCAGATGGGACTTA	1200
DB	698	GGCAGCTGAACACGACAGACCATGAAACCCGCTGGCGGCTGTCCAAATGGGCTGA	757
QY	1201	TTTATGTCATCTCAGAGCCCGGTGGAAAAACAGATCCTCTGGCTTCGCGAAAAAGATA	1260
DB	758	TCTATGTAACCCGGAAGGTGTTAACGGCCACCTGATCCGCTGAGAACCGCACGACG	817
QY	1261	TCAGGAAGCTTTTTCACGTATGGCCATGATGATGAGGAGACTGTGGCCCTGATCCCGG	1320
DB	818	TACTTGAACCTTCGCGCCCTATGGCGATGAACGACGAAAAAACCGAGCCCTCACAGCTG	877
QY	1321	GAGGGCATACATTTGTFAAAGACATGGTGCACGCTCTCTGAAAAATGTTTGGCGCAG	1380
DB	878	GGCGCCACACCGTCGGTAATGTACGGTAATGGCAATGCTCTGGCT---TAGCCCCG	934
QY	1381	GGCCTGATGGTCACCTGTGGAGGACGGGACTGGGATGGAAAAATAAATGTGGTACAG	1440
DB	935	ACCCAAAAGCCTCTGACGTTGAAAACACAGGCTTAGTGGGGCAACCCCAACATGCAG	994
QY	1441	GAACGGCAAAATATACCATCACCAAGTGGCCTGGAAGAGCCTGGTGCAC	1499
DB	995	GCAAGCAACCAACGCGTGACCTCGGGTATCGAAGTGTCTGGACAC	1043

RESULT 20

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 3.2%; Score 48; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 5.6e-05;
Matches 12; Conservative 217; Mismatches 157; Indels 0; Gaps 0;
QY 1065 GGATTTCGTGGGAGAGAGATGACTGGGAGTCGGACCTGGTATCTACTGGGGCCCTGAC 1124
DB 1445 GAAATTTGGTACRTRRR 1386
QY 1125 AACAGCCTCTGCAGATAACCGGATAAACCGGAACTTCAGAACTCTTGGCGCC 1184
DB 1385 RRR 1326
QY 1185 ACGCAGATGGGACTATTATGTCAATCTCAAGCCCGCTGAAACAGATCCCTCTG 1244
DB 1325 RRR 1266
QY 1245 GCTCCCGGAAGATATCAGGAAGACTTTTACGTATGGCCATGGATGATGAGGAGCT 1304
DB 1265 RRR 1206
QY 1305 GTGGCCCTGATCGGGGAGGCATACATTGGTAAACACATGCTGACGCTCTCTGAA 1364
DB 1205 RRR 1146
QY 1365 AAATGATTTGGCGCAGGCGCTGATGGTGCACCTGTGAGGAGCAGGAGCTGGGATGAAA 1424

Db 1145 RRR 1086
QY 1425 AATAAATGTTGTACAGGAACGGCAA 1450
DB 1085 RRR 1060

RESULT 21
US-09-125-287-2
; Sequence 2, Application US/09125287B
; Patent No. 6114602
; GENERAL INFORMATION:
; APPLICANT: BARG, Rivka
; APPLICANT: SALTS, Tehiam
; TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAOPY IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: INTRO GENETIC PARTHENOCAOPY IN PLANTS
; CURRENT APPLICATION NUMBER: US/09/125,287B
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: PCT/IL97/00051
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 4518
; TYPE: DNA
; ORGANISM: TPRP-F1 PROMOTOR
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6)..(191)
; OTHER INFORMATION: "n"s are any nucleic residue
US-09-125-287-2

Query Match 2.6%; Score 38.8; DB 3; Length 4518;
Best Local Similarity 43.8%; Pred. No. 0.045;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;
QY 424 CTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTGTCATCAATA 483
DB 3826 CTTATTGATGATTGTGGCAAGTATCCACCACCAAGACTTCAAGTGTCTTCAACTAATC 3885
QY 484 AAGTTGACTTTTGTATACACATCGGAATTCCTTAATCCGGAGCTATTCGATGATAA 543
DB 3886 AAGGTTTCCACTTTTCTCACTTTCATATTACTCCTACTCAATTTATGTGTGTAC 3945
QY 544 AAAAAACTCTCTCTGATCTTCTGCGCTATCGGGAGCTTTTCTACCGCTGTAG 603
DB 3946 AGTTGACATTTCAAGTATTAGGCCCAATTTCTAGCTCGGAATTTTAAATCTCTTT 4005
QY 604 CCGCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTTAACTCCTC 663
DB 4006 AAATATTTTGTATTATCTACTTATACGTAGTTTTCATAAGTATAAATTTCAATTCATA 4065
QY 664 TGAGATTACACACCCCTGAATCAATCCCTCGGGGCTGATTTGATTTGCCACAGAT 723
DB 4066 TATGAATTCACGCTCAAAAATTTTAAAGTTTATTAAACCAATGCCACATAACTTAGAACAT 4125
QY 724 TTCAACAGCTGGATATGGAGGCTCTCAAAAAGATATCAAGATTTTGTGACACTCCC 783
DB 4126 ACAATATCTTTTGTATCAAGATTTGGCAATTCGTATACAATAATCTTTAGCAAGTAATA 4185
QY 784 AGGATTGGTCCCTCGGGGATTTATGTT 809
DB 4186 TGTATACCAACATATTGTAATGAT 4211

RESULT 22
US-09-125-287-1
; Sequence 1, Application US/09125287B
; Patent No. 6114602
; GENERAL INFORMATION:
; APPLICANT: BARG, Rivka

APPLICANT: SALTS, Tehiam
TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAOPY IN
FILE REFERENCE: INTRIO GENETIC PARTHENOCAOPY IN PLANTS
CURRENT APPLICATION NUMBER: US/09/125,287B
CURRENT FILING DATE: 1998-11-09
EARLIER APPLICATION NUMBER: PCT/IL97/00051
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 12839
TYPE: DNA
ORGANISM: TPRP-F1 GENOMIC CLONE
FEATURE:
NAME/KEY: unsure
LOCATION: (5)..(11322)
OTHER INFORMATION: "n"s are any nucleic residue
US-09-125-287-1

Query Match 2.6%; Score 38.8; DB 3; Length 12839;
Best Local Similarity 43.8%; Pred. No. 0.083;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 424 CTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTGCAATCAATA 483
DB 3827 CTTATTGATGATTTGTGGCAAGTATCCACCAAGACTTCAAGTGTCTTCAACCTAAATC 3886
QY 484 AAGTTCATTTTCTATACAACTGCGAATTTCCCTTAATCCGGAGCTATTTCGATGATAA 543
DB 3887 AAGTTTCCACTTTTCTCACTTTCAATTAATTAATCACTCACTCACTCAATTAATGTTGATC 3946
QY 544 AAAAACTCTCTCTGTCTGTCT 603
DB 3947 AGTTGACATTTCAAGTATTAGGCCCAATTTTCTTAGCTCGAATTTTTTTTAAATCTCTTT 4006
QY 604 CCCTGATATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTCCCTC 663
DB 4007 AATATTTTGAATTTATATCTACTTTTATACGTATTTTATACGATATAAATTTTCAAT 4066
QY 664 TGAGATTACACAGCCCTGAATCAATCAATCCCTGGGGGCTGATTTTGAATTTATGCCACCAT 723
DB 4067 TATGAATTCACGTCATAAATTTTAAAGTTTATTAACTAATGCCATATAGAACAT 4126
QY 724 TTCAACAGCTGGATATGGAGCTCTGAAAGATATCAAGATTTGCTGACAACTTCCC 783
DB 4127 ACAATATCTTTTGTATCAAGATTTGCAATTCGTATACAAATATCTTTAGCAAGTAATA 4186
QY 784 AGATTGGTGGCTCGCGGATTTATGTT 809
DB 4187 TGTATACCAACATTATGTAATATGAT 4212

RESULT 23
US-08-991-789A-147/c
Sequence 147, Application US/08991789A
Patent No. 6225054
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
ATTORNEY/AGENT INFORMATION:
Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 292
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 147:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 base pairs

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A
FILING DATE: 11-Dec-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 210121.419C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 147:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 147:
US-08-991-789A-147

Query Match 2.3%; Score 33.8; DB 4; Length 579;
Best Local Similarity 57.8%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1316 CGCGGGAGGAGCATACATTTGGTAAAGCACATGTCGAGCGTCTCCTGAAAAATGTTATGG 1375
DB 427 CGGTGGAGCCATGAGTGGCACACGACGGGGAGCTGGTCTGTCAGAACGGCTTTG 368
QY 1376 CGCAGGGCCCTGATGGTGCACCTGTGGAGGAGCAGGAGGACTGGG 1417
DB 367 CTGACATCATCCGTTGACACAGCAAGGTGGAGGAGCTGG 326

RESULT 24
US-08-991-451-147/c
Sequence 147, Application US/09062451
Patent No. 6344550
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
ATTORNEY/AGENT INFORMATION:
Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 297
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 147:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-062-451-147

Query Match      2.3%; Score 33.8; DB 4; Length 579;
Best Local Similarity 57.8%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1316 CGCGGGAGGCGATACATTTGGTAAAGCAGCATGTGTCAGCGTCTCCTGAAAAATGTATTGG 1375
    || ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 427 CGGTGGAGCCAGTGAGATGGCACACGACGCGGCGAGCTGTCNTGTGCAAGCGCTTTG 368

Qy 1376 CGCAGGCGCTGATGTGTCACCTGTGTGGAGGAGCAGGACTGGG 1417
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 367 CTGACATCATCACCGTGTACACGACGAGAGGTGGAGGACGTGG 326

RESULT 25
US-09-598-326-147/C
; Sequence 147, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
;              Smith, John M.
;              Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;                     TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.K.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 147:

US-09-598-326-147

Query Match      2.3%; Score 33.8; DB 4; Length 579;
Best Local Similarity 57.8%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1316 CGCGGGAGGCGATACATTTGGTAAAGCAGCATGTGTCAGCGTCTCCTGAAAAATGTATTGG 1375
    || ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 427 CGGTGGAGCCAGTGAGATGGCACACGACGCGGCGAGCTGTCNTGTGCAAGCGCTTTG 368

Qy 1376 CGCAGGCGCTGATGTGTCACCTGTGTGGAGGAGCAGGACTGGG 1417
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 367 CTGACATCATCACCGTGTACACGACGAGAGGTGGAGGACGTGG 326
```

```
RESULT 26
US-08-716-351A-5/C
; Sequence 5, Application US/08716351A
; Patent No. 6033905
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based
; TITLE OF INVENTION: Retroviral Vectors
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,351A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03784
; FILING DATE: 06-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-128-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..10970
; OTHER INFORMATION: /standard_name= "p537 retroviral
; vector"
US-08-716-351A-5

Query Match      2.3%; Score 33.8; DB 3; Length 10970;
Best Local Similarity 51.7%; Pred. No. 3.3;
Matches 77; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 1267 AAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGCCCTGTGATCGCGGGAGGC 1326
    || ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 6662 AAGGTTTCTCAGATGGTAATAGATGCGGGGGATCAGCTGGACCTGGACACAGAAGTCT 6603

Qy 1327 ATACATTTGTTAAAGCACATGGTGACGCTCTCCTGAAAAATGTATTGGCGGAGGCCCTG 1386
    || || ||||| || || ||||| || || ||||| || || ||||| || || ||||| || ||
Db 6602 TTAGACTGATTAATAAACTAGGTGGAGAGCGAGGGGTGAGGCCACTGTCTGCGAGGCCAC 6543

Qy 1387 ATGGTGCACCTGTGGAGGAGCAGGACTG 1415
    || || ||||| || || ||||| || || ||||| || || ||||| || || ||||| || ||
Db 6542 CAGCTATGTTTTCAGGGGAGCAGACTG 6514

RESULT 27
US-09-134-001C-2054
; Sequence 2054, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
```



```
;
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-227C-7

Query Match      2.2%; Score 32.2; DB 2; Length 44377;
Best Local Similarity 53.6%; Pred. No. 26;
Matches 67; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 167 GGAAGCATCAGAAGGCGGGGAACTCCGTCGGCCAGTGAACCGTGCACACTCCGGG 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31666 GAGATCCCAAGAGGCGGGGTGGAACCGTGCGGCCCTGCACGAAGCGCCCTCGCGG 31607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 CAGTACATCCCGCGCGCTGATACCGCAAGAATGTCGCAAACTCCCGCTCCGTGCAG 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31606 CAGTACGAGTCCCGGCGCTGTCGGGGTCCGGGCGAGTACCTTCCCTCCACGTCACGCG 31547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 CGGGC 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31546 CGGTC 31542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 33
US-08-804-198-1/c
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kubstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostek, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
```

```
;
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-198-1

Query Match      2.2%; Score 32.2; DB 2; Length 44377;
Best Local Similarity 53.6%; Pred. No. 26;
Matches 67; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 167 GGAAGCATCAGAAGGCGGGGAACTCCGTCGGCCAGTGAACCGTGCACACTCCGGG 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31666 GAGATCCCAAGAGGCGGGGTGGAACCGTGCGGCCCTGCACGAAGCGCCCTCGCGG 31607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 CAGTACATCCCGCGCGCTGATACCGCAAGAATGTCGCAAACTCCCGCTCCGTGCAG 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31606 CAGTACGAGTCCCGGCGCTGTCGGGGTCCGGGCGAGTACCTTCCCTCCACGTCACGCG 31547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 CGGGC 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31546 CGGTC 31542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 34
US-09-156-316-11/c
; Sequence 11, Application US/09156316
; Patent No. 6183961
; GENERAL INFORMATION:
; APPLICANT: Bernstein, Harold S.
; APPLICANT: Coughlin, Shaun R.
; TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle
; FILE REFERENCE: UCSF-020/01US
; CURRENT APPLICATION NUMBER: US/09/156,316
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/060,688
; EARLIER FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2837
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-156-316-11

Query Match      2.1%; Score 32; DB 4; Length 2837;
Best Local Similarity 50.0%; Pred. No. 5.7;
Matches 106; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

QY 423 CCTAGTATTTTGTGCTGTAGTATCTATCCAGCAATAGGTATATCTCTTGCATCA-A 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2581 CATCAGTTTCAGCCTTCTAGAGTATGAAACCGGCAATTAATTAATCTCTGACAGATA 2522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 TAAAGTTGACTTTTGTATACAACATCGAATTTCCCTTAATCCGGAGCTATTCTGTATGAT 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2521 TAACTGTACTTCAGATTTTGTACTTAAGTCTCTTCTCCAGCAGCAATCAGCATAT 2462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 AAAAAAACTCTTCCTGTTCTGATCTCTTGGCGCTATCGGGAGGCTTTTCTACCGCTGT 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2461 CTATGTTGAAGTTCCCTTTCTCTTCTGTTGTCGCTGAACGCTCTTCTTTTAGACACTCT 2402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 602 AGCCGCTGATAAAAAGAGACTCAAAATTTCT 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 2401 AGCTCCGGGAATAGCAGAATCTTCATGTTT 2370

RESULT 35

US-09-442-055-1

; Sequence 1, Application US/09442055
; Patent No. 6306631
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Ehrlich, Stanislas Dusko
; TITLE OF INVENTION: No. 6306631el ppGpp Synthetase and Expression
; FILE OF INVENTION: Systems for Improved Production of Protein of
; FILE REFERENCE: 5259.204-US
; CURRENT APPLICATION NUMBER: US/09/442,055
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 0726/97
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2278
; TYPE: DNA
; ORGANISM: relA-bac
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(2222)
US-09-442-055-1

Query Match 2.1%; Score 31.8; DB 4; Length 2278;
Best Local Similarity 57.6%; Pred. No. 5.8;
Matches 57; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 292 TATTTCAGGATACCTTGTGTCATCAACACGCTACAAACAGACAGCTTTTGTCT 351
DB 618 TATTACAGATTTGCAACCTCATGAAGAAGAACGTCGACAGAGCTTTATGTCAT 677
QY 352 GACATCCCAAGAGAGGGAATATTCAGTCTGCCACCA 390
DB 678 GAGGTGTGAAGTGAAGAAGCGTGTGGAAGAAGTA 716

RESULT 36

US-09-282-147-38

; Sequence 38, Application US/09282147
; Patent No. 6274147
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram
; APPLICANT: YAO, Kun
; TITLE OF INVENTION: METHOD FOR GENERATING NONPATHOGENIC, INFECTIOUS
; TITLE OF INVENTION: PANCREATIC NECROSIS VIRUS (IPNV) FROM SYNTHETIC RNA
; FILE OF INVENTION: TRANSCRIPTS
; FILE REFERENCE: 8288-9023
; CURRENT APPLICATION NUMBER: US/09/282,147
; CURRENT FILING DATE: 1999-03-31
; EARLIER APPLICATION NUMBER: US/60/080,278
; EARLIER FILING DATE: 1998-03-31
; EARLIER APPLICATION NUMBER: PCT/US97/12955
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Infectious pancreatic necrosis virus
US-09-282-147-38

Query Match 2.1%; Score 31.8; DB 4; Length 3097;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1050 GGATTTAAACCGTGGGATTTGTCGCGGAAGAGAAGATGACTGGAGTCGGACCTGGTA 1109
DB 2060 GGAGTTCCAAAGGATGTTATTCACCGCGAAATCGCGGATGACGAGGAGACAATACCA 2119

QY 1110 TACTGGGGCCTGACACAAAGCCTCTTGCAG 1140
DB 2120 ATTCGGGTGTAGACATCAAGCCATCGCAG 2150

RESULT 37

US-09-196-390-5
; Sequence 5, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum L.
; STRAIN: cv. Florida
; TISSUE TYPE: ca. 21 d Caryopses
; IMMEDIATE SOURCE:
; LIBRARY: cDNA library in pBluescript sk (-)
; CLONE: pTASS1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 162...2559
US-09-196-390-5

Query Match 2.1%; Score 31.6; DB 4; Length 2825;
Best Local Similarity 55.5%; Pred. No. 7.7;
Matches 61; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

	QY	1212	CCTGAGGCCCGGTGGAAAACCAAGATCCTCTCTGGCTTCGCCGAAGAAGATATCATCGGGAAGCT	1271
	Db	918	CTGCAGCCCCCGCTGTACAAGAAGACCCTTTGGGATTTCAAGAATAATACATTTGGTTCGAG	977
	QY	1272	TTTTACGCTATGCCCATGGATGATGAGGAGACTGTGTGCCCTGTATCGCGGG	1321
	Db	978	GAGCCCGTGAGGCCAAGGATGATGCCCGGCGCTGTGCGAGATCATCGGG	1027

RESULT 38

US-08-365-486A-18/c
Sequence 18, Application US/08365486A
Patent No. 5834306
GENERAL INFORMATION:
APPLICANT: Webster, Keith A.
APPLICANT: Bishopric, Nanette H.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
TITLE OF INVENTION: Therapeutic Constructs
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,486A
FILING DATE: 23-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4353 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human NOS-1 gene, Fujisawa, et al,
INDIVIDUAL ISOLATE: J. Neurochem 63:140 1994
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4305
US-08-365-486A-18

Query Match 2.1%; Score 31.6; DB 2; Length 4353;
Best Local Similarity 54.2%; Pred. No. 10;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

[illegible]

Db
2852 TCTCCACACAGAAGACATCACAGGCTGCCTTGAAGACCTTCTTGGCCCCAGGTCTTGA 2795

RESULT 39

US-08-880-342-18/c
; Sequence 18, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; APPLICANT: Murphy, Brian
; APPLICANT: Laderoute, Keith R.
; APPLICANT: Green, Christopher J.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,342
; FILING DATE: 23-JUN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB95/00996
; FILING DATE: 13-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/365,486
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4353 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human NOS-1 gene, Fujisawa, et al,
; INDIVIDUAL ISOLATE: J. Neurochem 63:140 1994
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4305
US-08-880-342-18

```
Query Match      2.1%; Score 31.6; DB 4; Length 4353;
Best Local Similarity 54.2%; Pred. No. 10;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
```

[illegible]

Db 2852 TCTCCACACAGACATCACAGGCTGCTTGAAGACCTTCTTGCCCGAGGCTCTGA 2795

RESULT 40

US-08-365-486A-20/c
; Sequence 20, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,486A
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4780 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human NOS-SN gene, Nakane, et al,
; INDIVIDUAL ISOLATE: FEBS Lett 316:175 (1993)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 431..4732
US-08-365-486A-20

Query Match 2.1%; Score 31.6; DB 2; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1270 CTTTTCACGTATGCCATGATCATGAGGACTGTGCCCTGATCGCGGAGGGCATA 1329
Db 3339 CTCCTCCAGCTGCGATCATCTGATGAGGAAATTTGGCCCTTTTCAATGTTGACATCA 3280
QY 1330 CATTGGTAAAGCACATGGTGCAGCGTCTCTGAAAAATGATTGGCGCAGGCGCTGA 1387
Db 3279 TCTCCACACAGACATCACAGGCTGCTTGAAGACCTTCTTGCCCGAGGCTCTGA 3222

RESULT 41

US-09-123-708-3/c
; Sequence 3, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GODECKE, Axel

; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2003
; CURRENT APPLICATION NUMBER: US/09/123,708
; CURRENT FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/553,503
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: P4411402.8
; EARLIER FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4780
; TYPE: DNA
; ORGANISM: Cytomegalovirus
US-09-123-708-3

Query Match 2.1%; Score 31.6; DB 3; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1270 CTTTTCACGTATGCCATGATCATGAGGACTGTGCCCTGATCGCGGAGGGCATA 1329
Db 3339 CTCCTCCAGCTGCGATCATCTGATGAGGAAATTTGGCCCTTTTCAATGTTGACATCA 3280
QY 1330 CATTGGTAAAGCACATGGTGCAGCGTCTCTGAAAAATGATTGGCGCAGGCGCTGA 1387
Db 3279 TCTCCACACAGACATCACAGGCTGCTTGAAGACCTTCTTGCCCGAGGCTCTGA 3222

RESULT 42

US-09-123-624-3/c
; Sequence 3, Application US/09123624
; Patent No. 6149936
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2004
; CURRENT APPLICATION NUMBER: US/09/123,624
; CURRENT FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 08/553,503
; PRIOR FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: 4411402.8
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4780
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-123-624-3

Query Match 2.1%; Score 31.6; DB 3; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1270 CTTTTCACGTATGCCATGATCATGAGGACTGTGCCCTGATCGCGGAGGGCATA 1329
Db 3339 CTCCTCCAGCTGCGATCATCTGATGAGGAAATTTGGCCCTTTTCAATGTTGACATCA 3280
QY 1330 CATTGGTAAAGCACATGGTGCAGCGTCTCTGAAAAATGATTGGCGCAGGCGCTGA 1387
Db 3279 TCTCCACACACAGACATCACAGGCTGCTTGAAGACCTTCTTGCCCGAGGCTCTGA 3222

RESULT 43

US-08-880-342-20/c
; Sequence 20, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.

APPLICANT: Bishopric, Nanette H.
APPLICANT: Murphy, Brian
APPLICANT: Laderoute, Keith R.
APPLICANT: Green, Christopher J.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/880,342
FILING DATE: 23-JUN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB95/00996
FILING DATE: 13-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/365,486
FILING DATE: 23-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 4780 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human NOS-SN gene, Nakane, et al,
INDIVIDUAL ISOLATE: FEBS Lett 316:175 (1993)
FEATURE:
NAME/KEY: CDS
LOCATION: 431..4732
US-08-880-342-20

Query Match 2.1%; Score 31.6; DB 4; Length 4780;
Best Local Similarity 54.2%; Pred. No. 11;
Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 1270 CTTTTCACGTATGGCCATGATGAGGAGACTGTGGCCCTGATCGGGAGGGCATA 1329
Db 3339 CTCTCCAGTCGCATATTGCTGATGAGGGAATTTGGCCCTTTCAATGTTGACATCA 3280
Qy 1330 CATTGGTAAAGACATGGTGCAGCGTCTCCTGAAAATGATATGGCGAGGGCCTGA 1387
Db 3279 TCCTCCACAGAAGACATCACAGGCTGCCTTGAAGACCTTCTTGGCCCGAGGTCCTGA 3222

RESULT 44
US-08-253-155A-7
Sequence 7, Application US/08253155A
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jenio
APPLICANT: Draetta, Giulio

TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-028
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3407 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-253-155A-7

Query Match 2.1%; Score 31.4; DB 1; Length 3407;
Best Local Similarity 51.8%; Pred. No. 10;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 114 TTCTCCAGCCAGTTAAGCCAGCGGTTTCCTGCTGAAAAATGTCGGCAAAACGGGGAAGC 173
Db 1059 TTCCTAAGCGGCTGTACAAGGCCCTCTCCCTGCTGAAAGAGGAATTTGAACTGAGCAAGC 1118
Qy 174 ATCAGAGAGGGGGGGGAACTCCGTCGGCCAGTGAACCGTCCACACTCCGGGCGAGTACA 233
Db 1119 TGCAGCAGCGCTGGGGCGGAGGTGGAGAGAAGATCAAGCAGACCCCGTAAGTACC 1178
Qy 234 TCGCGCGCGCGCTGATA 250
Db 1179 TGCTGCAGGAGCAGCTA 1195

RESULT 45
US-09-007-005-17
Sequence 17, Application US/09007005B
Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rihe
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
CURRENT FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
EARLIER FILING DATE: 1997-11-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA


```

: CURRENT FILING DATE: 1999-07-20
: EARLIER APPLICATION NUMBER: PCT/EP98/00388
: EARLIER FILING DATE: 1998-01-19
: NUMBER OF SEQ ID NOS: 9
: SEQ ID NO 5
: LENGTH: 4160
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: MAKAR17
: LOCATION: 945..968
: OTHER INFORMATION: region corresponding to
: FEATURE:
: NAME/KEY:
: LOCATION: 1444..1445
: OTHER INFORMATION: insertion point of T-DNA
: OTHER INFORMATION: insertion in the line
: FEATURE:
: NAME/KEY:
: LOCATION: 1418..1570
: OTHER INFORMATION: putative open reading
: US-09-341-678-5

```

	Query Match	2.1%	Score 30.8;	DB 4;	Length 4160;
	Best Local Similarity	55.7%;	Pred. No. 18;		
	Matches 59;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;
QY	736	ATATGAGGCTGTGAAAAGATATCAAGATTTGCTGACAACCTCCAGGATTGTGCC	795		
Db	4035	ATATCTTGGCTTTGGAAATGAAGTGCAGATTTATGCAGACAATGGTTCAGGTTTGACCG	3976		
QY	796	CTGCGGATTATGCTCATATTATGGTCTCTTTTATTCGTATGGCTTG	841		
Db	3975	AAACTGACGTGTTTAAAAACCTTCATATATATTTTTGTATCTCTG	3930		

RESULT 49

```

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

	Query Match	2.1%	Score 30.8	DB 4	Length 4403765
	Best Local Similarity	57.1%	Pred. No. 67		
	Matches	56	Conservative 0	Mismatches 42	Indels 0
					Gaps 0
QY	887	CAGTGTGGTCTAGCAACAGTTT	TGAACCGGTGAACACTGCGGGATTAACGT	TAATCTGA	946
Db	4113218	CAGTACGGTGGCGACGATCG	GGCCGTACAACATCTGACGGGATCCGGCTATTGGGC	4113277	
QY	947	TAAAGCCGTCGATTCTGTGGCCAGTCAAGAAAAAT	984		
Db	4113278	ACAAGCCAGTTCCTATTGGGCACAGCCCAATTAGAAT	4113315		

```

RESULT 50
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      2.1%; Score 30.8; DB 4; Length 4411529;
Best Local Similarity 57.1%; Pred. No. 67;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      887  CAGTGGTGGTCAGCAACGTTTTCGAACCGCTGGAACAGCTGGCCGGGATAACGTTAATCTGGA 946
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      4121023  CAGTGACGGTCTGCGGACGATCGCGGGCGGTACACATCTGACCGGGATCCGGCTATTGGGGC 4121082

QY      947  TAAAGCCCGTCGATTGCTGTGGCCAGTCACAGAAAAAT 984
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      4121083  ACAAGCCAGTTCCTATTGGGCACAAGCCAAATTAGAAAT 4121120

```

Search completed: June 19, 2003, 03:36:03
Job time : 91.1139 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 01:20:59 ; Search time 1574.88 Seconds

(without alignments)
15312.305 Million cell updates/sec.

Title: US-09-674-277-1

Perfect score: 1489

Sequence: 1 ctgcagtcggagatgaaag.....ctggaaggagcctggtgcac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	197.2	13.2	762	14	BQ752115
2	196	13.2	707	14	BQ751801
3	191	12.8	639	14	BQ751512
4	185.2	12.4	742	14	BQ751881
5	183.2	12.3	606	14	BQ751387
6	176.8	11.9	729	10	AW179968
					AW179968 MGA0020F

7	176.4	11.8	562	12	BG370112
8	171.8	11.5	737	17	CNS01HHH
9	161	10.8	573	10	AW180683
10	159.4	10.7	465	12	BG278355
11	154.6	10.4	546	14	BQ142683
12	152.6	10.2	786	17	BH403595
13	134.2	9.0	571	10	AW180185
14	132.6	8.9	600	10	AW180270
15	104.4	7.0	474	14	BH865076
16	97	6.5	417	14	BH869928
17	87.6	5.9	718	13	BI750328
18	82.8	5.6	291	14	BH868464
19	79.2	5.3	278	14	BH862783
20	66	4.4	540	14	BH038121
21	66	4.4	540	14	BH038189
22	55.8	3.7	396	10	AW792137
23	55.8	3.7	397	10	AW792143
24	53.2	3.6	399	12	BF739529
25	50.2	3.4	634	12	BG810098
26	47.8	3.2	591	13	BI874487
27	47	3.2	877	9	AI068496
28	45.8	3.1	537	12	BF251526
29	44.8	3.0	660	13	BI718013
30	43.8	2.9	335	10	AW789645
31	41.4	2.8	640	13	BI717417
32	40.4	2.7	1101	17	CNS001N6
33	40	2.7	694	14	BQ539839
34	39.8	2.7	543	17	BH119903
35	39.6	2.7	661	12	BG711981
36	39.4	2.6	775	12	BF494150
37	39.4	2.6	803	17	CNS018FX
38	38.2	2.6	717	14	C90436
39	38.2	2.6	1095	17	CNS00AS8
40	38.2	2.5	599	10	AW707196
41	37.8	2.5	635	10	BE556340
42	37.8	2.5	652	9	AI980795
43	37.8	2.5	627	10	AW180016
44	37.6	2.5	630	17	DR22018S
45	37.6	2.5	587	14	BQ611766
46	37.2	2.5	476	12	BF940954
47	37	2.5	520	10	AW981658
48	37	2.5	637	17	BH762769
49	37	2.5	1024	17	CNS005BT
50	37	2.5	211	9	AA601988
51	36.8	2.5	587	10	AW065188
52	36.8	2.5	876	17	CNS01NE5
53	36.8	2.5	626	14	BQ819458
54	36.6	2.5	694	17	AZ090819
55	36.6	2.5	922	17	CNS015Y8
56	36.6	2.4	475	17	B49229
57	36.4	2.4	514	12	BF495230
58	36.4	2.4	529	17	AQ850214
59	36.4	2.4	344	12	BG629704
60	36.2	2.4	454	10	AW638187
61	36.2	2.4	706	12	BF035003
62	36.2	2.4	928	17	AG135931
63	36.2	2.4	279	9	AV072812
64	36	2.4	702	17	AZ978404
65	36	2.4	835	9	AJ454978
66	36	2.4	967	12	BG387766
67	36	2.4	633	14	BQ390739
68	35.8	2.4	589	12	BG322881
69	35.6	2.4	787	17	CNS010B7
70	35.6	2.4	1051	17	CNS014GC
71	35.6	2.4	1101	17	CNS0100X
72	35.6	2.4	506	17	AQ514856
73	35.4	2.4	715	17	AQ259871
74	35.4	2.4	1118	12	BE874869
75	35.4	2.4	2321	11	AY108097
76	35.4	2.4	412	14	W09301
77	35.2	2.4	538	17	AZ723610
78	35.2	2.4	565	17	BH401331
79	35.2	2.4			

BG370112	LMA123 V8
AL144438	Anopheles
AW180683	MGA0831f
BG278355	a3g03np.r
BQ142683	113 Metar
BH403595	AG-ND-127
AW180185	MGA0268f
AW180270	MGA0361f
BH865076	mgap008xD
BH869928	mgns008XE
BI750328	FG02_0740
BH868464	ngcs014XP
BH862783	mgcm004XM
BH038121	FucusSS12
BH038189	FucusSS22
AW792137	D01025-R
AW792143	D01029-R
BF739529	NCM10B4T3
BG810098	ngct002xn
BI874487	963115H08
AI068496	mgae0002c
BF251526	EST418787
BI718013	103102300
AW789645	C01253-F
BI717417	103101960
AL061217	Drosophill
BQ539839	PTAM0150
BH119903	RPCI-24-3
BG711981	pg11n.pk0
BF494150	AT02578_5
AL109287	Drosophill
C90436	C90436 Dict
BH724397	BOMBA22TF
AL056173	Drosophill
AW07196	sk22a06.y
BE556340	sq01d10.y
AI980795	pat.pk003
AW180016	MGA0072f
AL745989	banlo rer
BQ611766	sap65a03.
BF940954	hu62a05.x
AW981658	PC15B10.P
BH762769	BMBAC330H
AL058090	Drosophill
AA601988	no88h06.s
AW065188	ST40G03.P
AL152110	Anopheles
BQ819458	1030078A0
AZ090819	RPCI-23-4
AL106058	Drosophill
B49229	RPCI11-5114
BF495230	AT04106.5
AQ850214	LMAJFV1.1
BG629704	cc-esf1cL
AW638187	bl67B04.W
BF035003	6014359082
AG135931	Fan trogl
AV072812	AV072812
AZ978404	2M0254B16
AJ454978	AJ454978
BG387766	602412644
BQ390739	NISC_mq14
BG322881	EMI_15_CO
AL098749	Drosophill
AL104118	Drosophill
AL098379	Drosophill
AQ514856	hb_5208_B
AQ259871	nbxb0024H
BE874869	601488919
AY108097	Zea mays
W09301	ma63d03.r1
AZ723610	RPCI-24-6
BH401331	AG-ND-137

[illegible]

```

QY 812 TTATGCTCTTCTTTTATCGTATGCTTGGCAGCTGCGGAGACATACAGGACATATGA 871
Db 332 CTACGGCGGTCTCTTATCGCATGGCTTGGCAGCGCGGACGCTACCGAGTTCCAGCA 391
QY 872 TGGCCGGGAGGCGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 931
Db 392 CGGACCGGAGGTGGTGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 451
QY 932 TAACGTTAATCTGGATAAGACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 991
Db 452 CAATGTCTGCTGGGAGGCGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 511
QY 992 CAGTATTTCTGGGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051
Db 512 CAAGATCTCTGGGCGGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
QY 1052 ATTAAACAGCTGGGATTTGCTGGCGGAGAGAGAGATGATGCTGGGAGTGGGAGCTGGTATA 1111
Db 572 TTTCCAGAGCGCGGCTTCTCCGAGGCGGCTCCGACACCTGGGAGCGGACGAGTCCGT 631
QY 1112 CTGGGGCGCTGACAACAA 1129
Db 632 CTACTGGGGCGCGAGAA 649

```

RESULT 2

```

BQ751801
LOCUS EST632364 DSCT Colletotrichum trifolii cDNA clone pDSCT7-68, mRNA
DEFINITION

```

```

ACCESSION BQ751801.1 GI:21907206
VERSION EST.
KEYWORDS Colletotrichum trifolii.
SOURCE Colletotrichum trifolii.
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.

```

REFERENCE

```

AUTHORS Samad,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F., and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)
COMMENT Contact: Deborah A. Samad
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@umn.edu
TIGR sequence name: MTSAG68TV More information is available at:
www.medicago.org
Seq primer: (gta ata cga ctc act ala ggg c).

```

FEATURES

source

```

1. 707
/organism="Colletotrichum trifolii"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT7-68"
/clone_lib="PSCN"
/tissue_type="mycelia"
/dev_stage="young, actively growing mycelia (3 days after
inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)."
/lab_host="DH5alpha"
/note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
EcoRI; isolate: 2sp2; cDNA was prepared from polyA+
enriched RNA The cDNA was ligated into lambda gtl1 from
Stratagene and packaged using Gigapack packaging extracts.
An aliquot of the amplified library was used to transduce
E. coli 11090 and phage DNA was purified from a liquid
lysate. The cDNA inserts were gel purified after EcoRI

```

digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

BASE COUNT 146 a 247 c 195 g 119 t

Query Match 13.2%; Score 196; DB 14; Length 707;

Best Local Similarity 62.0%; Pred. No. le-50;

Matches 310; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

```

QY 632 CTACTATCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTCAATCAATCC 691
Db 178 CTGGTGGCGGATGCGCTGAAGCTCACATCTCCGCCACACACGCCCTCACCACCC 237
QY 692 CTGGGGGGCTGATTGATTATGCCACAGATTTCACACAGCTGGATATGAGGCTCTGAA 751
Db 238 GCTGGCGCAGGACTTTGACTTACGTCGCGCGCTTCAAGTCTCGACTACGAGGCGTCAA 297
QY 752 AAAAGATATCAAGATTGCTGACAACTTCCAGGATTGTCGCCCTGCGGATTATGTCA 811
Db 298 GAAGGACCTCAGCGGCTGATGACCGACTCCAGGACTGGTGGCTGCCGACTTTGGCCA 357
QY 812 TTATGTCCTTCTTTATTCGTATGCTTGGCACGTTGGCGGAGACATACAGGACATATGA 871
Db 358 CTACGGCGGTCTGTTATCCGATGGCTGGCACAGCGCGGACGCTACCGAGTTCCAGCA 417
QY 872 TGCCCGGGGAGGCGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 931
Db 418 CGGACCGGAGGTGGTGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 477
QY 932 TAACGTTAATCTGGATAAGACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 991
Db 478 CAATGTCTGCTGGGAGGCGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 537
QY 992 CAGTATTTCTGGGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051
Db 538 CAAGATCTCTGGGCGGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
QY 1052 ATTAAACAGCTGGGATTTGCTGGCGGAGAGAGATGATGCTGGGAGTGGGAGCTGGTATA 1111
Db 598 GTTCCAGAGCGCGGCTTCTCCGAGGCGGCTCCGACACCTGGGAGCGGACGAGTCCGT 657
QY 1112 CTGGGGCGCTGACAACAGC 1131
Db 658 CTACTGGGGCGCGAGAAC 677

```

RESULT 3

BQ751512

LOCUS

DEFINITION BQ751512 639 bp mRNA linear EST 18-JUL-2002

EST632075 DSCT Colletotrichum trifolii cDNA clone pDSCT5-66, mRNA

sequence.

ACCESSION BQ751512

VERSION BQ751512.1 GI:21906917

KEYWORDS EST.

SOURCE Colletotrichum trifolii.

ORGANISM Colletotrichum trifolii

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;

mitosporic Phyllachoraceae; Colletotrichum.

REFERENCE 1 (bases 1 to 639)

AUTHORS Samad,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,

Cheung,F., and Fraser,C.M.

TITLE ESTs from mycelia of Colletotrichum trifolii race 1

JOURNAL Unpublished (2002)

COMMENT Other_ESTS: EST632074

Contact: Deborah A. Samad

Department of Plant Pathology

University of Minnesota

495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA

Tel: 612 625 1243

Fax: 651 649 5058

Email: debbyspuccini.crl.umn.edu
 TIGR sequence name: MTSAE66TV More information is available at:
 www.medicago.org
 Seq primer: (gta ata cga ctc act ata ggg c).
 Location/Qualifiers
 1..639
 /organism="Colletotrichum trifolii"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST5-66"
 /clone_lib="DST"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."
 BASE COUNT 131 a 222 c 176 g 110 t
 ORIGIN

Query Match 12.8%; Score 191; DB 14; Length 639;
 Best Local Similarity 64.2%; Pred. No. 3.8e-49;
 Matches 287; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
 QY 632 CTACTATCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 DB 182 CTGGTGGCGGATGCGGTGAGCTACATCTCTCCGCCAGCACAGCCGCTCACCACCC 241
 QY 692 CTGGGGGCGTATTTGATTATGACACAGATTTCAACAGCTGGATGAGGCTCTGAA 751
 DB 242 GCTGGCCAGAGACTTTGACTAGTCGCGGCTTCAAGTCGCTCAGCTACGAGGCGCTCA 301
 QY 752 AAAGATATCAAGATTTCCTACAACTTCCACAGATTGTGCGCTCGGATATGTCA 811
 DB 302 GAAGACCTCAGCCCTGATGACCGACTCCAGACTGTGGCTCTGCCACTTTGGCCA 361
 QY 812 TTATGGTCTTTCTTTATTCGTATGCTTGGCAGGTGCGGACATACAGGACATGA 871
 DB 362 CTACGGCGGTCTGTATCCGCATGCTGSCACAGCGCGGACGTACCGAGTTACGA 421
 QY 872 TGGCGGGAGGCGCCAGTGTGTGACACAGTTTGAACCGGTGAACAGCTGGCGGA 931
 DB 422 CGGACGGGAGGTGGTGAGAGGCGCCAGCAACGCTTCGCCGCTCAACAGCTGGCGGA 481
 QY 932 TAAGCTTAATCTGATAAGCCCTCGATTGCTGTGCGCAGTCAAGAAAATACGGCTC 991
 DB 482 CAATGTAGCTTCACAGGCCCTCGCTGCTGTGCGCCATCAAGAAAATACGGCAA 541
 QY 992 CAGTATTTCTGGGAGACGTGATGCTCTGACTGTAATCTTGCCCTTGAATCCATGGG 1051
 DB 542 CAAGATCTCGTGGCCGACCTCATGATCTTGGCGGCAAGTGGCCCTCAGTCCATGGG 601
 QY 1052 ATTTAAACGCTGGGATTTGCTGGCG 1078
 DB 602 TTCCAGACGGGCGGCTTCTCCGAGG 628

RESULT 4
 B0751881
 LOCUS B0751881 742 bp mRNA linear EST 18-JUL-2002
 DEFINITION DSCT Colletotrichum trifolii cDNA clone pDST8-27, mRNA
 sequence.
 accession B0751881

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

B0751881.1 GI:21907286
 EST
 Colletotrichum trifolii.
 Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 1 (bases 1 to 742)
 Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
 Cheung,F. and Fraser,C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST632443
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbyspuccini.crl.umn.edu
 TIGR sequence name: MTSAE66TV More information is available at:
 www.medicago.org
 Seq primer: (gta ata cga ctc act ata ggg c).
 Location/Qualifiers
 1..742

FEATURES
 source

/organism="Colletotrichum trifolii"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST8-27"
 /clone_lib="DST"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."
 BASE COUNT 151 a 259 c 212 g 120 t
 ORIGIN

Query Match 12.4%; Score 185.2; DB 14; Length 742;
 Best Local Similarity 62.0%; Pred. No. 2.9e-47;
 Matches 310; Conservative 0; Mismatches 188; Indels 2; Gaps 1;
 QY 632 CTACTATCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 DB 178 CTGGTGGCGGATGCGGTGAGCTACATCTCTCCGCCAGCACAGCCGCTCACCACCC 237
 QY 692 CTGGGGGCGTATTTGATTATGCCACAGATTTCACAGCTGGATATGAGGCTCTGAA 751
 DB 238 GCTGGCCAGAGACTTTGACTAGTCGCCGCTCTCAAGTCGCTCAGCTACGAGGCGGTCAA 297
 QY 752 AAAAGATATCAAGATTTCCTGACAACTTCCAGGATTGTCGCCCTGGCGGATTTATGGTCA 811
 DB 298 GAAGGACCTCAGCGGCTGATGACCGACTCCCGAGGACTGGTGGCTCTGGCGGCTTGGCCA 357
 QY 812 TTATGGTCTTTCTTTATTCGTATGCTTGGCAGCGGTGGCGGAAATACAGGACATATGA 871
 DB 358 CTACGGCGGTCTGTTCATCCGATGCGCTGGCACAGCGCGGACGCTACCGAGTTACGA 417
 QY 872 TGGCGGGGGA--GGCGCCAGTGTGTGACGAAACGTTTGAACCCCTGAACAGCTGGCGG 929
 DB 418 CGGACGGGAGGGGTGGGAGAGGGCGGACGAAACGCTTCGCAACCCCTCAACAGCTGGCGG 477

/organism="Anopheles gambiae"

/strain="PEST"

/db_xref="taxon:7165"

/clone="08M05"

/clone_lib="NotreDamel"

/note="end : Sp6"

BASE COUNT 161 a 217 c 224 g 127 t 8 others
ORIGIN

Query Match 11.5%; Score 171.8; DB 17; Length 737;

Best Local Similarity 57.3%; Pred. No. 5.7e-43;

Matches 357; Conservative 6; Mismatches 253; Indels 7; Gaps 3;

QY 852 GGAACATACAGGACATATGATGCCCGGGAGGCGCCAGTGGTGCAGCAACGTTTTGAA 911

DB 12 GTACCATATCGCTCCCTGGAYGGCGCGTGGGCGAGGACGCGCTCAACAGCGCTTG-- 69

QY 912 CCCTCAACAGCTGGCGGATACGTTAATCTGGATAAAGCCGCTCGATTGCTGTGGCCA 971

DB 70 CCACTGAATCTCCGCGGATACGTAAGCTGGATAAAGCCGCTCGATTGCTGTGGCCT 129

QY 972 GTCAGAAAATAATACGGCTCCAGTATTTCTTGGGAGACCTGATGCTCTGACTGTAAT 1031

DB 130 ATCAAGAAAATAATACGCAAAAATCTCTGGCGGACCTGTTATCTCGGGTAAAC 189

QY 1032 GTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGAAGAGATGAC 1091

DB 190 GTGGCGCTGAGAACTCCGCTCTCCGTAACCTTTGGTTTGGTGGCGGCTGAAGACGTC 249

QY 1092 TGGGACTCGGACCTGTATCTACTGGGCGCTGACAAAGCCCTTTCCAGATAACCGGAT 1151

DB 250 TGGGAACCGGATCTTCACTGTAACCTGGGTGATGAATAAACCTGGCTGAMCCACCGGAC 309

QY 1152 AAAAAGGGAAATCTCAGAAACCTCTTGGCGCACGACATGGGACTTATTTATGTAAT 1211

DB 310 CCGGAAGCGTGCTAARCTCTCTGGCAGCCACCGAATGGGCTGATCTACGTGAAC 369

QY 1212 CTTGAAGCCCGGTGAAACACGATCTCTGGTTCGCGAAGATATCAGGAAGCT 1271

DB 370 CCGGAAGG--CCGAAGCCGCGGTGAACCGCTGTGAGCGCGGACGATTCGTGGGACC 427

QY 1272 TTTTACGATATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGATACA 1331

DB 428 TTCCGCAACATGGGATGACGATGAGAGACCGCTCGCGTGAATGACGCGGCCACAC 487

QY 1332 TTTGGTAAGCACATGGTGCAGCGCTCTCTGMAAATATGTTGGCGAGGCGCTGATGT 1391

DB 488 CTGCGCAAGACTCACGG--CGCGGTGAGCCACCCACCTGGTACCGACCCGAGCC 544

QY 1392 GCACCTGTGGAGGAGGAGGACTGGGATGGAATAAATATGTTGGTACAGGAACGCAAA 1451

DB 545 TCACCGATTAAAMRCACARGCGCTGGGCTGGGCGGACGACGCGGCGGTGGCGCT 604

QY 1452 TATACCATACAGTGGCGCTGGA 1474

DB 605 GATGCCATTACCTCCGCTGGA 627

RESULT 9

AW180683

LOCUS

DEFINITION Mga Library Mycosphaerella graminicola cDNA clone MGA0831

5' similar to catalase-peroxidase, mRNA sequence.

AW180683

VERSION

AW180683.1 GI:6447878

KEYWORDS

SOURCE

ORGANISM

Mycosphaerella graminicola.

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et

Chaetothromycetes Incertae sedis; Mycosphaerellaceae;

Mycosphaerella.

REFERENCE

1 (bases 1 to 573)

Authors: Keon, J.P.R., Bailey, A.M. and Hargreaves, J.A.

TITLE

JOURNAL

MEDLINE

COMMENT

DEFINITION AG-ND-127D17.TF ND-TAM Anopheles gambiae genomic clone AG-ND-127D17
DNA sequence.
ACCESSION BH403595
VERSION BH403595.1 GI:17349811
KEYWORDS GSS.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE 1 (bases 1 to 786)
Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
Direct Submission of BAC-end sequences from Anopheles gambiae
Unpublished (2001)
Other_GSSs: AG-ND-127D17.TR
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..786
/organism="Anopheles gambiae"
/strain="PEST"
/db_xref="taxon:7165"
/clone="AG-ND-127D17"
/clone_lib="ND-TAM"
/note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT 183 a 211 c 148 g 244 t
ORIGIN

Query Match 10.28; Score 152.6; DB 17; Length 786;
Best Local Similarity 62.09; Pred. No. 8.2e-37;
Matches 298; Conservative 0; Mismatches 169; Indels 14; Gaps 3;
QY 1019 CCTGACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAACGC-TGGGATTTGCTGGCG 1077
DB 729 CCTAACCGGTAAAGCTTGGCTTGGAACTATGGGATTAACAACCAAGGATATTTGGCGGTG 670
QY 1078 GAAGAGAAGATGACTGGGACCTGGACCTGGTATACTGGGGCGCTGACAAAGCCTCTTG 1137
DB 669 CCGGTACAGATGATGAAGACCGATAAAGATGTATCTTGGGAGATGAGAAACCTGG 610
QY 1138 CAGATAACCGGGATAAAA-----ACGGGAAACTTCAGAAACCTCTTGGCGCACG 1187
DB 609 CTGGGAACAGATAAAGATATACAGGGGATCGTGATCTGGAAGATCGCTGGCAGCAACA 550
QY 1188 CAGATGGGACTTATTATGCAATCTCGAAGGCCCGCGTGGAAACACAGATCTCTGGCT 1247
DB 549 ACAATGGGGCTTATCTATGTAATCTCTGAAGGACCGGAAGAAACCGGATCGGCTGGCA 490
QY 1248 TCCGCGAAGATATCAGGGAAGCTTTTTCACGTATGCCATGGATGATGAGGACACTGTG 1307
DB 489 GCTGCTAAAGATATCCGTGTATACATTTGGCCCGTATGGGATGAATGACGAGGAACCTGTA 430
QY 1308 GCCCTGATCGCGGAGGGCCATACATTTTGGTAAAGACACATGGTGACGGCTCTCCTGAAAAA 1367
DB 429 GCGCTAATTCGTGGTGGCCACACATTTGGTAAACTCATGTGGCCGAGATGACGCGCTA 370
QY 1368 TGTATTGGCAGGGCCTGATGTTGACCTGTGTGAGAGAGAGGACTGGGATGGAAAAAT 1427

DB 369 ---GTAGGAGCTGAGCCTGAGGATCGGAATTGAAGCTCAGGATTAGTGTGGCAAGT 313
QY 1428 AAATGTGTACAGGAACCGCAATATACCATCATCAGCTGCGCTGGAGGAGCCTGTGTCG 1487
DB 312 AAATTGGACAGCTCGGAGCAGATGCTATTACCAAGTGTCTGGGAAGTACCTGGACC 253
QY 1488 A 1488
DB 252 A 252
RESULT 13
AW180185
LOCUS
DEFINITION MGA0268f Mga Library Mycosphaerella graminicola cDNA clone MGA0268
5, similar to PEROXIDASE/CATALASE, mRNA sequence.
ACCESSION AW180185
VERSION AW180185.1 GI:6447380
KEYWORDS EST.
SOURCE Mycosphaerella graminicola.
ORGANISM Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyriomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella.
REFERENCE 1 (bases 1 to 571)
AUTHORS Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.
TITLE A group of expressed cDNA sequences from the wheat fungal leaf
blotch pathogen, Mycosphaerella graminicola (Septoria tritici)
JOURNAL Fungal Genet. Biol. 29 (2), 118-133 (2000)
MEDLINE 20374020
COMMENT Contact: Hargreaves JA
Cell Biology Department
IACR-Long Ashton Research Station
Long Ashton, Bristol, BS41 9AF, UK
Tel: +44(0)1275 392181
Fax: +44(0)1275 394281
Email: john.hargreaves@bbsrc.ac.uk
Insert Length: 2500 Std Error: 0.00
Seq primer: M13 reverse.
FEATURES
source
location/Qualifiers
1..571
/organism="Mycosphaerella graminicola"
/strain="Strit"
/db_xref="taxon:54734"
/clone="MGA0268"
/clone_lib="Mga Library"
/note="Vector: pSPORT1; Library constructed from cultures
utilizing ammonium ions as a source of nitrogen"
BASE COUNT 127 a 177 c 141 g 126 t
ORIGIN
Query Match 9.08; Score 134.2; DB 10; Length 571;
Best Local Similarity 61.08; Pred. No. 5.2e-31;
Matches 268; Conservative 0; Mismatches 168; Indels 3; Gaps 3;
QY 632 CTACTATCCAGAACACTGGATTTAACTCTCTGAGATTACACAGCCCTGAATCAAATCC 691
DB 131 CTGGTGGCCAAATGAGCTGACACCAAGATCTCGCCACGACACCCGCTGCTACCGACC 190
QY 692 CTGGGGGGCTGATTTTGTATTATGCCACAGATTTCACACAGC-TGGATATGGAGGCTCTGA 750
DB 191 ATTCGGCAACGAGTTCGACTACCCAGCGCATCAAGAAGCTTGATTAACAATGGCCCTCA 250
QY 751 AAAAGATATCAAGATTTGCTGACAACTTCCACAGATTGGTGGCCCTGCGGATATGGTC 810
DB 251 AGAAGGACTTGAACGACCTCATGACCGACTCGAAGGACTTCTGGCCACGACAGACTTTGGCC 310
QY 811 ATTATGTCCTCTTTTTCGTTATGCTTGGCAGCGTGGCCGACATACAGGACATATG 870
DB 311 ACTACGGTGGTTTCTTCGTCGGATGGCATGGCAGCGGTGTACCTACCGTGTGCGCTG 370
QY 871 ATGCCCGGGGAGGCGCAGTGGTGTGAGCAACCTTTTGAACCGCTGAACAGCTGGCGCG 930


```

Email: rdean@clmsonson.edu
Seq primer: T3 primer (AATTAAACCTCACTAAAGGG)
High quality sequence stop: 69.
Location/Qualifiers
1..877
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgae0002c01f"
/clone_lib="Magnaporthe grisea Appressorium Stage cDNA
Library"
/dev_stage="Germinated conidia on appressorium-inductive
surface"
/note="Vector: pBluescriptII SK(+) Vector; Site1: EcoRI;
Site2: XhoI; The appressorium formation-specific cDNA
library was constructed from conidia germinated for 5-8
hr on an inductive surface. The library has an average
insert size of 1.5 kbp."
202 a 248 c 243 g 178 t 6 others
BASE COUNT
ORIGIN
Query Match 3.2%; Score 47; DB 9; Length 877;
Best Local Similarity 51.2%; Pred. No. 0.0053;
Matches 110; Conservative 0; Mismatches 105; Indels 0; Gaps 0
QY 801 GATTATGTCATTATGTCCTTCTTTATTCGTATGCTTGGCACGTCGCCGAACATAC 860
DB 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 |||||
29 GACGATGCTACTAGGCCCGCTCTCGTGGCTTGGCTTGGCACGTCAGCTACTAC 88
QY 861 AGGACATATGATGCCGGGAGGCCGACGTGGTGGTCAGCAACGCTTTTGAACGCTGAAC 920
DB 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 |||||
89 GACAAGGAGACCGCACTGGCGGCAGCAACGGTGCCACCATGCGCTTCTCGCCTGAGGCT 148
QY 921 AGCTGGCCGATACGTTATCTATCTGGATAAAGCCGCTCGATTGCTGGCCAGTCAAGAA 980
DB 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 |||||
149 GCGCACGCTGCCAACGCTGGTCTGAAGGCGCGCGTGACCTCTGGAGCCTATCAAGGCC 208
QY 981 AAATACGCTCCAGTATTTCTCTGGGAGACCTGAT 1015
DB 11 ||||| 11 ||||| 11 ||||| 11 ||||| 11 |||||
209 AAGTACCCTGGATCACTACTGACCTCTGGAT 243

RESULT 28
BF251526
LOCUS
DEFINITION
ES7418787 Coccidioides immitis spherule cDNA library Coccidioides
immitis cDNA clone C1A110 5' sequence, mRNA sequence.
BF251526
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Coccidioides immitis.
Coccidioides immitis.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.
1 (bases 1 to 537)
Gardner, M.J. and Kirkland, T.
Generation of ESTs from Coccidioides immitis spherule cDNA library
Unpublished (2000)
Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.
Location/Qualifiers
1..537
/organism="Coccidioides immitis"
/db_xref="taxon:5501"
/clone="C1A110"
/clone_lib="Coccidioides immitis spherule cDNA library"
/dev_stage="spherule"
/lab_host="S01R"

FEATURES
source

```


Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."		92 a	231 c	229 g	107 t	1 others
BASE COUNT						
ORIGIN						
Query Match	3.0%;	Score 44.8;	DB 13;	Length 650;		
Best Local Similarity	53.1%;	Pred. No. 0.023;				
Matches 94;	Conservative 0;	Mismatches 83;	Indels 0;	Gaps 0;		
QY	801	GATTATGGTCATTATGTCCTCTTTCTTATTCGATATGGCTTGGCAGCGTGGCGGACATAC	860			
Db	465	GAGCAGCGCAGCTACGGCCCCGCTGCTGGTGGCGCTGGCAGCGCTCCGGACCGTAC	524			
QY	861	AGGACATATATGTCGGGGAGGCGCCAGTGGTGTGTCAGCAACGTTTGAACCGCTCAAC	920			
Db	525	GCCNAGAGACGCGCAGCGGGCCAGCAACGGGGCCACCATTGGCTTTGCTCCGAGTGC	584			
QY	921	AGCTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCTGAATTGCTTGGCCCAAG	977			
Db	585	GAGTGGGGGCCCAACGCGGGTCTGGCGGTGCTCGCAAGCTGCTGGAGCCCGTCAAG	641			
RESULT 30						
AW789645						

DEFINITION	Co1253-F Lambda zap, stragatagene Blumeria graminis f. sp. hordei cDNA clone Co1253 similar to catalase, mRNA sequence.
ACCESSION	AW789645
VERSION	AW789645.1 GI:13901242
KEYWORDS	EST.
SOURCE	Blumeria graminis f. sp. hordei.
ORGANISM	Blumeria graminis f. sp. hordei
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Erysiphales; Erysiphaceae; Blumeria.
AUTHORS	1 (bases 1 to 335) Thomas,S.W., Rasmussen,S.W., Glaring,M.A., Rouster,J.A. and Oliver,R.P.
TITLE	Gene identification in the fungal pathogen Blumeria graminis by expressed sequence tag analysis
JOURNAL	Unpublished (2000)
COMMENT	Contact: Rasmussen,S.W. Department of Yeast Genetics Carlsberg Laboratory 10 Gl. Carlsbergvej, DK-2500, Copenhagen, Denmark Tel: 45 3327 5230 Fax: 45 3327 4766 Email: sw@erc.dk High quality sequence stop: 335 POLYA-No.
FEATURES	Location/Qualifiers 1..335 /organism="Blumeria graminis f. sp. hordei"

/organism="Blumeria graminis f. sp. hordei" /db_xref="taxon:62688" /clone="C01253" /clone_lib="Lambda Zap, Stratagene" /cell_type="conidia" /lab_host="Hordeum vulgare" /lab_host="Hordeum vulgare"		94 a 70 C 78 g 91 t 2 others	Query Match Best Local Similarity Matches	2.9%; Score 43.8; DB 10; Length 335; 66.3%; Pred. No. 0.031; 0; Mismatches 32; Indels 0; Gaps 0;	1067 63
QY	1008	GACCTGATGTCCTGACTGGTAATGTCCTTCGAATCCAGGATTTAAACCGCTGGA	1067		
Db	4	GATCTTTACCTATTACTGGTAATGTCGATTTGAATCAATGGTGTGAACACCTTTGGA	63		
OY	1068	TTTGCTGGCGGAAGAAGATGACTGGGAGTCGGA	1102		

Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pjejong@gmail.com). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 316 row: M column: 7
 Seq primer: SP6
 Class: BAC ends.

FEATURES	Location/Qualifiers
source	1..543
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="RPCI-24-316M7"
	/clone_lib="RPCI-24"
	/sex="Male"
	/cell_type="Spleen/Brain"
	/note="vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT	203 a	76 c	94 g	170 t
ORIGIN				
Query Match	2.7%	Score 39.8;	DB 17;	Length 543;
Best Local Similarity	50.8%;	Pred. No. 0.8;		
Matches	95;	Conservative	0;	Mismatches 92; Indels 0; Gaps 0
QY	590	TTCTACCGCTGTAGCCGCTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACT	649	
Db	423	TTGTACACAGACGATCCTGATACATGAAGATTCAAAATCAGTCAAACTCAGTAAAAATC	364	
QY	650	GGATTTAACTCCCTGAGATTACACAGCCCTGATCAATCCCTGGGGGGCTGATTTGA	709	
Db	363	AGGGTCAAAATATTTTGTGTTTTCTGCTCAATTACCACAGTCTTTGGTTTACATTTTCTGAA	304	
QY	710	TTATGCCACCAGATTTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTT	769	
Db	303	ATTTAGACAGATGTTCCCACTGGACATATAGTTTATATATAAAGTCTCTGATTTT	244	
QY	770	GCTGACA	776	
Db	243	ATTTCGA	237	

RESULT	35
BG711981	
LOCUS	661 bp mRNA linear EST 08-MAY-2001
DEFINITION	pBln.pk010.g23: Normalized Liver Library Gallus gallus cDNA clone pgln.pk010.g23 5' similar to dbj BAB14712.1 (AK023886) unnamed protein product [Homo sapiens]G, mRNA sequence.
ACCESSION	BG711981
VERSION	BG711981.1 GI:14005931
KEYWORDS	EST.
SOURCE	chicken.
ORGANISM	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
REFERENCE	Phasianinae; Gallus. 1 (bases 1 to 661)
AUTHORS	Burnside,J., Morgan,R.W. and Cogburn,L.A.
TITLE	Chicken Ests from a normalized liver library
JOURNAL	Unpublished (2001)
COMMENT	Contact: Joan Burnside Molecular Endocrinology University of Delaware

```

/organism="Drosophila melanogaster"
/db xref="taxon:7227"

```

Query Map
Best Loc:

Query Match 2.6%; Score 39.4; DB 17; Length 803;
Best Local Similarity 19.0%; Pred. NO. 1.4;


```

QY 1059 ACCTGGGATTTCTGCGGGAAGAGAGACTGCGGAGTC 1099
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 ACTATTAACTTTGTCCTCCGGGAGAAAAGATTCAATGGAATC 409

RESULT 42
BE556340
LOCUS
DEFINITION
  sq01d10.y1 Gm-cl045 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
  Gm-cl045-1868 5' similar to TR:Q39780 Q39780 ASCORBATE PEROXIDASE.
  ; mRNA sequence.
ACCESSION
  BE556340
VERSION
  BE556340.1 GI:9820830
KEYWORDS
  EST.
SOURCE
  soybean.
ORGANISM
  Glycine max
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 635)
  Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corvelli,V., Khanna
  A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
  Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr
  R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
  R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: estewatson.wustl.edu
  This clone is available through: ResGen, Invitrogen Corp. 2130
  South Memorial Parkway Huntville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: ccuresgen.com
  Insert Length: 1062 Std Error: 0.00
  High quality sequence stop: 426.
FEATURES
  source
    1..635
    /organism="Glycine max"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-cl045-1868"
    /clone_lib="Gm-cl045"
    /tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
    /lab_host="DH10B"
    /note="Vector: pBluescriptII SK+; Site_1: EcoRI; Site_2:
    XhoI; This cDNA library was constructed from mRNA isolated
    from etiolated hypocotyl tissue of 9-10 day old seedlings
    of the cultivar Williams 82. Complementary DNA was
    synthesized from mRNA using a primer consisting of a
    poly(GT) primer with a XhoI restriction site. EcoRI
    adapters were ligated to the blunt-ended cDNA fragments
    followed by digestion with EcoRI and XhoI. The cDNA
    fragments were directionally cloned into the EcoRI-XhoI
    restriction site of the pBluescript vector. The ligated
    cDNA fragments were transformed into DH10B host cells
    (Gibco BRL). This library was constructed by Dr. Randy
    Shoemaker."
  BASE COUNT 196 a 137 c 146 g 155 t 1 others
  ORIGIN
    Query Match 2.5%; Score 37.8; DB 10; Length 635;
    Best Local Similarity 50.9%; Pred. No. 3.8;
    Matches 143; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
    QY 819 CTTTTCCTTTTCTGCTGGCGGGAACATACAGGACATATGATGGCGG 878
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 150 CCTCTCATGCTTCGATTAGCTGGCATGATGCTGGTACATTACCATGCTAAACAACACA 209

QY 879 GGAGGGCCCGCTGGTGTCTAGCAACAGTTTGTGAACCGCTGAACAGCTGGCGCGGATAACGTT 938
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 GGAGGGCCCTAATGTTCTATCAGAACACAGAGAGTTGAATCAC---GCAGCAACAACAG 266

QY 939 AATCTGTATAAGCCCGCTCGATTGCTGTGCCAGTCAAGAAAATAATACGCTCCAGTATT 998
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 GGACTGGAACACAGCACTTGCTTCTGTGAGGAAGTGAAGGCCAAACA---TCCAAAATTT 323

QY 999 TCCTGGGGAGACCTGATGTGCTCCAGCTGGTAATGTTGCCCTTGAATCCATGGGATTTAAA 1058
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 324 TCATATGCCGACCTTTACCAGCTAGCTGGTGTGTTGTCAGTAGAAGTTACCGGGGGGCCA 383

QY 1059 ACCTGGGATTTCTGCGGGAAGAGAGACTGCGGAGTC 1099
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 ACTATTAACTTTGTCCTCCGGGAGAAAAGATTCAATGGAATC 424

RESULT 43
AI980795
LOCUS
DEFINITION
  pat.pk0033.f8.f chicken activated T cell cDNA Gallus gallus cDNA
  clone pat.pk0033.f8.f 5' similar to amyloid precursor protein, mRNA
  sequence.
ACCESSION
  AI980795
VERSION
  AI980795.1 GI:5883823
KEYWORDS
  EST.
SOURCE
  chicken.
ORGANISM
  Gallus gallus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
  Phasianidae; Gallus.
REFERENCE
  1 (bases 1 to 652)
  Tirunagaru,V.G., Sofer,L., Cui,J. and Burnside,J.
  An expressed sequence tag database of T-cell-enriched activated
  chicken splenocytes: sequence analysis of 5251 clones
  Genomics 66 (2), 144-151 (2000)
  20318616
  Contact: Joan Burnside
  Molecular Endocrinology
  University of Delaware
  40 Townsend Hall, Newark, DE 19717, USA
  Tel: 302 831-1345
  Fax: 302-831-3411
  Email: joan@udel.edu, www.chickest.udel.edu
  Seq primer: T7.
FEATURES
  source
    1..652
    /organism="Gallus gallus"
    /db_xref="taxon:9031"
    /clone="pat.pk0033.f8.f"
    /clone_lib="Chicken activated T cell cDNA"
    /sex="male"
    /cell_type="Con A-activated splenic T cell"
    /lab_host="E.coli TOP10 F'"
    /note="Vector: pcdNA3"
  BASE COUNT 186 a 124 c 165 g 166 t 11 others
  ORIGIN
    Query Match 2.5%; Score 37.8; DB 9; Length 652;
    Best Local Similarity 49.7%; Pred. No. 3.9;
    Matches 96; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
    QY 1210 ATCTGAAGGCCCGGTGGAAAACACAGATCCTCTGGCTTCGGGAAAGATATCAGGGAAG 1269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 2 AACCTAGAATCAGCTATGGAATCATGCTCTCATGCTTCTATGCACGGAACACTAAGACCA 61

QY 1270 CTTTTTCACATATGGCCATCGATGATGAGGAGACTGTGGCCCTGATCGGGAGGGGATA 1329
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CAGTTGAACCTTCTTCCAGTGGATGGAGAGTTCAGTCTCGATGACCTTCAGCCATGGCATC 121

QY 1330 CATTTGGTAAGGCACATGGTGCAGCGCTCTCTGAAATAATGTATTTGGCGGAGGCGCTGATG 1389
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      122 CCTTGTGTGATCTGTTCCAGCCAACTACTGAAATGAAGTTGAGCCTGTTGATGCC 181
QY      1390 GTGCACCTGTGGA 1402
Db      182 GCCCAGCTGCAGA 194

RESULT 44
LOCUS   AW180016
DEFINITION  MGA0072f Mga Library Mycosphaerella graminicola cDNA clone MGA0072
5', similar to catalase-peroxidase, mRNA sequence.
ACCESSION AW180016
VERSION   AW180016.1 GI:6447211
KEYWORDS  EST.
SOURCE    Mycosphaerella graminicola.
ORGANISM  Mycosphaerella graminicola.
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
          Chaetochyriomycetes incertae sedis; Mycosphaerellaceae;
          Mycosphaerella.
REFERENCE 1 (bases 1 to 627)
AUTHORS   Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.
TITLE     A group of expressed cDNA sequences from the wheat fungal leaf
          blotch pathogen, Mycosphaerella graminicola (Septoria tritici)
JOURNAL   Fungal Genet. Biol. 29 (2), 118-133 (2000)
MEDLINE   20374020
COMMENT   Contact: Hargreaves JA
          Cell Biology Department
          IACR-Long Ashton Research Station
          Long Ashton, Bristol, BS41 9AF, UK
          Tel: +44(0)1275 392181
          Fax: +44(0)1275 394281
          Email: john.hargreaves@bbsrc.ac.uk
          Insert Length: 2500 Std Error: 0.00
          Seq primer: M13 reverse.

FEATURES             Location/Qualifiers
     source           1..627
                     /organism="Mycosphaerella graminicola"
                     /strain="Strit"
                     /db_xref="taxon:54734"
                     /clone="MGA0072"
                     /clone_lib="Mga Library"
                     /note="Vector: pSPORT1; Library constructed from cultures
                     utilizing ammonium ions as a source of nitrogen"

BASE COUNT      146 a 204 c 160 g 117 t
ORIGIN

Query Match      2.5%; Score 37.6; DB 10; Length 627;
Best Local Similarity 59.3%; Pred. No. 4.4;
Matches 64; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY      1382 GCCTGATGGTGCACCTGTGGAGGAGCAGGACTGGGATGGGAAAAATAAATGTGGTACAGG 1441
Db      1 GCCAACGGCGCCCTCCATCGAGGAGCAGGAGGTCTCGGATGGAGAACAGCAGTACAGGACGG 60

QY      1442 AACCGGCAATATACCATCACCAGTGCCTGGCAGGAGGAGCCTGGTCGAC 1489
Db      61 CAAGGGTCTTAACCCCATCATCATCCGGATTGGAGTCATCTGGACGTC 108

RESULT 45
LOCUS   DR22018S/C
DEFINITION  Danio rerio genomic clone dkey-22018, genomic survey sequence.
ACCESSION AL745989
VERSION   AL745989.1 GI:21349287
KEYWORDS  GSS.
SOURCE    zebrafish.
ORGANISM  Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 630)

```

```

AUTHORS   Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE     Direct Submission
JOURNAL   Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
          Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
          humquery@sanger.ac.uk Unpublished
COMMENT   This sequence was generated from the SP6 end of BAC 22018. 22018 is
          part of the Danokey Pilot BAC Library created by R. Plasterk and
          N.V. Keygene.
          Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES             Location/Qualifiers
     source           1..630
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="Dkey-22018"
                     /tissue_type="Testis"
                     /note="vector pindigobAC-536"

BASE COUNT      259 a 59 c 82 g 230 t
ORIGIN

Query Match      2.5%; Score 37.6; DB 17; Length 630;
Best Local Similarity 47.1%; Pred. No. 4.4;
Matches 115; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY      446 ATCTATCCCGCAGCAATAGGTATATATCTCTGTCATCATATAAGTTGATGATACACA 505
Db      630 ATATATACATGGAGTATGAGTATCATATTTGTTGTTACTTTTAAATTCAGCAGCTATA 571

QY      506 TCGGAATTTCCCTTAATCCGGAGCTATTCGTATGATATAAAAAACCTTCCTGTTCTGAT 565
Db      570 CACTTATATATTTATATATACACTTATTCATCATATATTAATTAATGATGCTCATTT 511

QY      566 TCTTCTGGCGCTATCCGGGAGCTTTTCTACCGCTGTAGCGGTGATAAAAAGAGACTCA 625
Db      510 TCTTCTTGACATTCAGATGATCTTTTTTTTGTCATTAACACATCATTAATTAATAGC 451

QY      626 AAATTTCTACTATCCAGAAACACTGGATTTAACTCTCTGAGATTACACAGCCTCAATC 685
Db      450 ATTGATCATTTTAACTAATTAACACAGCTTTTACCATTAAATTAATTAACACATCATTA 391

QY      686 AAT 689
Db      390 TAAT 387

RESULT 46
LOCUS   BQ611766
DEFINITION  sap65a03.v1 Gm-cl087 Glycine max cDNA clone SOYBEAN CLONE ID:
          Gm-cl087-5382 5' similar to TR:Q9SXL5 Q9SXL5 CHLOROPLAST ASCORBATE
          PEROXIDASE PRECURSOR ;, mRNA sequence.
ACCESSION BQ611766
VERSION   BQ611766.1 GI:21601435
KEYWORDS  EST.
SOURCE    soybean.
ORGANISM  Glycine max
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
          Glycine.
REFERENCE 1 (bases 1 to 587)
AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvett,V., Khanna
          ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
          Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
          ,Y., Person,B., Swallier,T., Gibbons,M., Pape,D., Harvey,N., Schurk
          ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
          ,R., Waterston,R. and Wilson,R.
          Public Soybean EST Project
          Unpublished (1999)
          Contact: Shoemaker R/Public Soybean EST Project
          Public Soybean EST Project
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800

```


AUTHORS NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE
JOURNAL

TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)

Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
Sequenced from the Brugia malayi BAC library constructed by Claire Whiffon and Dr Mike Quail. The sequence was generated by The Wellcome Sanger Institute, The Sanger Institute, Cambridge, UK in collaboration with Mark Blaxter, ICAPB, University of Edinburgh, Edinburgh, UK.
seq primer: T7 (TAATACGACCTCACTATAGGG)

Email: mark.blaxter@ed.ac.uk
Sequenced from the Brugia malayi BAC library constructed by Claire
Whitton and Dr Mike Quail. The sequence was generated by The
Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in
collaboration with Mark Blaxter, ICAPB, University of Edinburgh,
Edinburgh, UK.
Seq primer: T7 (TAATACGACTCACTATAGGG)
Class: BAC ends.
Location/Qualifiers
1..637
/organism="Brugia malayi"
/strain="TRS"
/db_xref="taxon:6279"
/clone_lib="Brugia malayi Genomic BAC Library 3"
/sex="Mixed (male and female)"
/tissue_type="whole parasite"
/dev_stage="microfilaria (L1)"
/note="vector: pBAC3.6; Site 1: BamH I; Brugia malayi
genomic DNA was partially cleaved with Sau3A I and size
fractionated. 7,392 clones were generated with mean insert
size ~48 kbp. The library was constructed by Claire
Whitton, Blaxter Nematode Genetics Lab, University of
Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing
Unit, The Sanger Centre, Cambridge, UK."

genomic DNA was partially cleaved with Sau3A I and size fractionated. 7 392 clones were generated with mean insert size ~48 kbp. The library was constructed by Claire Whitton, Blaxter Nematode Genetics Lab, University of Edinburgh, UK, and Dr Mike Quail, the Pathogen Sequencing Unit, The Sanger Centre, Cambridge, UK.

186 a 100 c 121 g 230 t
Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing Unit, The Sanger Centre, Cambridge, UK."

h 2.5%; Score 37; DB 17; Length 637;
Similarity 50.9%; Pred. No. 6.9;

Similarity 50.9%; Score 37; DB 17; Length 637;
88; Conservative 0; Mismatches 85; Indels 0; Gaps 0

7 AGAGACTCAAAATTTCTACTATATCCAGAAACACTGGATTAACTCCTCTGAGATTACACAG 676

4 ACAGACTAAATAATTTTGCAACCATAAAGCTTGAATAATAAAATGACAAAATCACAGCA 265

4 ACAGACTAAATAATTTTGCAACCACTAAACTTGAATAATAAAATGACAAAAATCACAGCA 265

7 CCTGAATCAAAATCCTGGGGGGCTGATTTTGATTATGCCACCAGATTTC AACACGCTGGA 736

7 CCCTGAATCAAATCCCTGGGGGCTGATTTTGATTATGCCACCAGATTTCAACAGCTGGA 736
4 CTATTAGATCATACCTTCAGATCATTTTGAGAAATCAACCACTTGTGAATAAATTCC 205

4 CTATTAGATCATACCTTCAGATCACTTTGAGAAATCAACCACTTGTGAATAATCC 205

7 TATGGAGGCTCTGAAAAAAGATATCAAAGATTTGCTGACAACATCCCAAGGAT 789

[illegible]

4 TGTGGATTGTCTGCAAAACGACATAAAAATCTCGTCGAAGCTTCCATTCATT 152

CNS005BT	1024 bp	DNA	linear	GSS 03-JUN-1999
Proserchia melanocarpa				

CNS005BT 1024 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR11006 of RPCI-98 library from *Drosophila melanogaster* (fruit
fly)

BACR11006 of RPCI-98 library from *Drosophila melanogaster* (fruit fly), genomic survey sequence.
AL058090

AL058090
AL058090.1 GI:4932671
GSS.

GSS.
Drosophila melanogaster.
Drosophila melanogaster

Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera. Brachycera: Muscomorpha.

1. Eukaryota; Metazoa; Artthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
(phases 1 to 1024)

Epiphyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 1024)
Genoscope.
Direct Submissions

Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

- Web : www.genoscope.cns.fr

Search completed: June 19, 2003, 03:33:42
Job time : 1591.88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 13:22:38 ; Search time 250.397 Seconds
(without alignments)
13391.616 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagtcgagatgaaag.....ctggaaggagcctggtcgac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : N_Geneseq_101002:*

Result No.	Score	Query Match	Length	DB ID	Description
1	1489	100.0	1489	21	AAZ36101 Nucleic acid seque
2	357.6	24.0	2238	19	AAV06555 Microscilla furves
3	298.4	20.0	2235	19	AAV06555 Mycobacterium tube
4	298.4	20.0	2331	18	AAV04000 Mycobacterium tube
5	298.4	20.0	2331	20	AAV00817 M. tuberculosis ca
6	298.4	20.0	4403765	22	AAI99683 Mycobacterium tube
7	298.4	20.0	4411529	20	AAI99682 Mycobacterium tube
8	296.8	19.9	2331	22	AAV00818 M. tuberculosis ca
9	274.2	18.4	2221	19	AAV89968 Mycobacterium tube

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1489	100.0	1489	21	AAZ36101 Nucleic acid seque
2	357.6	24.0	2238	19	AAV06555 Microscilla furves
3	298.4	20.0	2235	19	AAV06555 Mycobacterium tube
4	298.4	20.0	2331	18	AAV04000 Mycobacterium tube
5	298.4	20.0	2331	20	AAV00817 M. tuberculosis ca
6	298.4	20.0	4403765	22	AAI99683 Mycobacterium tube
7	298.4	20.0	4411529	20	AAI99682 Mycobacterium tube
8	296.8	19.9	2331	22	AAV00818 M. tuberculosis ca
9	274.2	18.4	2221	19	AAV89968 Mycobacterium tube

10	274.2	18.4	4795	14	AAQ51531 M.tuberculosis H37
11	274.2	18.4	4795	20	AAV16969 M.tuberculosis kat
12	250.4	16.8	2196	11	AAQ06815 Sequence encoding
13	224.6	15.1	2262	19	AAV06554 Alcalligenes (Deley
14	122.8	8.2	660	21	AAA49864 Mycobacterium tube
15	95.2	6.4	1852	21	AAV07490 Fusarium venenatum
16	91.4	6.1	538	24	ABQ23234 Oligonucleotide fo
17	91.4	6.1	538	24	ABQ23235 Oligonucleotide fo
18	75.6	5.1	150	22	AAV95160 katG gene probe #2
19	66.8	4.5	150	22	AAV95161 katG gene probe #3
20	63.8	4.3	150	22	AAV95159 katG gene probe #1
21	47.4	3.2	538	24	ABQ23236 Oligonucleotide fo
22	47.4	3.2	538	24	ABQ23237 Oligonucleotide fo
23	46.4	3.1	150	22	AAV95162 katG gene probe #4
24	38.8	2.6	1358	21	AAV11909 Aspergillus oryzae
25	38.8	2.6	4320	18	AAV89278 TRP-P1 gene, incl
26	37.8	2.5	5591	23	ABL08713 Drosophila melanog
27	37.8	2.5	6000	23	ABL08655 Drosophila melanog
28	36.6	2.5	14752	20	AAV07256 Borrelia burgdorfe
29	36.4	2.4	801	21	AAV07750 Fusarium venenatum
30	35	2.4	327	22	AAH53165 S. epidermidis ope
31	35	2.4	3803	22	AAH54197 S. epidermidis gen
32	34.8	2.3	824	23	AAV73010 DNA encoding novel
33	34.8	2.3	6240	23	AAV90138 DNA encoding novel
34	34.6	2.3	526	22	AAK55269 Human immune/haema
35	34.6	2.3	6065	24	AAV70580 Chemically treated
36	34.6	2.3	6065	24	AAV61261 Human gene regulat
37	34.6	2.3	6065	24	ABK31357 Signal transductio
38	34.4	2.3	544	20	AAZ42098 Human endometrium
39	34.2	2.3	800	20	AAV15175 Human gene express
40	34.2	2.3	893	23	AAV49611 Staphylococcus aur
41	34.2	2.3	927	23	AAV55352 Staphylococcus aur
42	34.2	2.3	1164	23	AAV52250 Staphylococcus aur
43	34.2	2.3	2554	18	AAV74876 Human cDNA fragmen
44	34	2.3	188	24	AAV94435 Human foetal cDNA,
45	34	2.3	308	22	AAH94492 Human foetal cDNA,
46	34	2.3	316	22	AAV93991 DNA encoding novel
47	34	2.3	636	22	AAV30450 Human reproductive
48	34	2.3	636	22	AAV06230 Maize nitrite redu
49	34	2.3	1844	12	AAV12000 Human immune/haema
50	34	2.3	16605	22	AAK65489 Human immune/haema
51	34	2.3	16605	22	AAK87049 Human immune/haema
52	33.8	2.3	457	22	ABA44433 Human breast cell
53	33.8	2.3	457	22	ABA54879 Human foetal liver
54	33.8	2.3	457	22	ABA24644 Probe #3110 for ge
55	33.8	2.3	457	22	AAK03156 Human brain expres
56	33.8	2.3	457	22	AAK28607 Human bone marrow
57	33.8	2.3	579	18	AAV83336 Breast cancer tumo
58	33.8	2.3	579	19	AAV68878 DNA molecule encod
59	33.8	2.3	579	21	AAV80859 Human breast tumou
60	33.8	2.3	579	24	AAV46749 Human breast tumou
61	33.8	2.3	579	24	AAV99705 Breast tumour-spec
62	33.8	2.3	848	23	AAV85884 DNA encoding novel
63	33.8	2.3	848	23	AAV88775 DNA encoding novel
64	33.8	2.3	1086	23	AAV77622 DNA encoding novel
65	33.8	2.3	1086	23	AAV85860 DNA encoding novel
66	33.8	2.3	1086	23	AAV88715 DNA encoding novel
67	33.8	2.3	4171	23	AAV77682 Chemically treated
68	33.8	2.3	6754	24	ABL70346 Human gene regulat
69	33.8	2.3	6754	24	AAV61305 Retro virus vector
70	33.8	2.3	10970	15	AAV73735 Human immune/haema
71	33.8	2.3	13496	22	AAV67371 Drosophila melanog
72	33.6	2.3	2620	23	ABL24238 DNA encoding novel
73	33.6	2.3	3165	23	AAV90422 Human differential
74	33.6	2.3	3176	22	AAV81799 Human bone marrow
75	33.6	2.3	3936	22	AAV90058 DNA encoding novel
76	33.6	2.3	4653	23	AAV73484 DNA encoding novel
77	33.6	2.3	4653	23	AAV89265 DNA encoding novel
78	33.6	2.3	4653	23	AAV89808 DNA encoding novel
79	33.6	2.3	4653	23	AAV90173 cDNA sequence enco
80	33.6	2.3	4845	22	AAV09227 Human neuroblastom
81	33.4	2.2	838	22	AAV94561 Staphylococcus epi
82	33.4	2.2	1074	24	AAV92591 Staphylococcus epi

83	33.4	2.2	4590	22	AAH24065	Yeast AOD9604-asso
84	33.4	2.2	6383	23	ABK42528	Genomic sequence #
85	33.4	2.2	21732	23	ABAI4492	Human nervous syst
86	33.4	2.2	21732	22	AAK81131	Human immune/haema
87	33.4	2.2	21732	23	ABK42527	Genomic sequence #
88	33.2	2.2	1003	22	AAH23806	Human transferase
89	33.2	2.2	1154	23	ABV23389	Human prostate exp
90	33.2	2.2	1154	23	ABV29245	Human prostate exp
91	33.2	2.2	3226	21	AAH77665	Human PRO1154 cdNA
92	33.2	2.2	3226	21	AAH77665	Human prostate bound pro
93	33.2	2.2	3226	21	AAH77665	Human PRO1154 (UNQ)
94	33.2	2.2	3226	22	AAH77665	Human prostate bound pro
95	33.2	2.2	3226	22	AAH77665	Human PRO1154 (UNQ)
96	33.2	2.2	4590	22	AAH24065	Yeast AOD9604-asso
97	33.2	2.2	4590	22	AAH24065	Human secreted pro
98	33.2	2.2	5459	22	AAH35012	Human colon cancer
99	33.2	2.2	5459	22	ABL90756	Human polynucleoti
100	33.2	2.2	6765	22	AAH90108	Human bone marrow
101	33	2.2	710	22	AAH03304	Human cDNA clone (
102	33	2.2	1026	22	AAH03304	Iron uptake ABC tr
103	33	2.2	1811	22	AAH15242	Human cDNA sequenc
104	33	2.2	27540	24	ABK83161	Pneumococcal patho
105	32.8	2.2	697	20	AAZ33534	Human prostate can
106	32.8	2.2	789	22	AAH81702	Human differentia
107	32.8	2.2	880	24	AAH47781	Hansenula polymorp
108	32.8	2.2	2149	21	AAH77736	Human cancer assoc
109	32.8	2.2	7086	22	AAH77415	Human nervous syst
110	32.8	2.2	7086	22	AAH03024	Human reproductive
111	32.8	2.2	7388	22	AAH17416	Human nervous syst
112	32.8	2.2	7388	22	AAH03025	Human reproductive
113	32.6	2.2	510	22	ABM48451	Human breast cell
114	32.6	2.2	510	22	ABA66355	Human foetal liver
115	32.6	2.2	510	22	ABA33417	Probe #11883 for g
116	32.6	2.2	510	22	AAK14772	Human brain expres
117	32.6	2.2	510	22	AAK40512	Human bone marrow
118	32.6	2.2	510	22	AAI46546	Probe #15232 used
119	32.6	2.2	567	22	AAH33403	Human immune/haema
120	32.6	2.2	570	22	AAH13369	Human cDNA clone (
121	32.6	2.2	2325	24	ABN79846	Fungal ZBC gene se
122	32.6	2.2	2408	23	ABL19631	Drosophila melanog
123	32.6	2.2	2739	23	ABL20333	Drosophila melanog
124	32.6	2.2	3073	22	AAH16293	Human cDNA sequenc
125	32.6	2.2	8881	23	ABL19630	Drosophila melanog
126	32.6	2.2	8881	23	ABL20332	Drosophila melanog
127	32.6	2.2	34980	22	AAH66431	Pyrococcus abyssi
128	32.4	2.2	460	23	ABV61149	Human prostate exp
129	32.4	2.2	486	22	AAK02037	Human brain expres
130	32.4	2.2	486	22	AAK27493	Human bone marrow
131	32.4	2.2	486	22	AAI33406	Probe #2092 used t
132	32.4	2.2	575	22	ABA63361	Human foetal liver
133	32.4	2.2	575	22	ABA76620	Human foetal liver
134	32.4	2.2	575	22	ABA30585	Probe #9031 for ge
135	32.4	2.2	575	22	ABA11128	Probe #19594 for g
136	32.4	2.2	575	22	AAK37604	Human bone marrow
137	32.4	2.2	575	22	AAK51262	Human bone marrow
138	32.4	2.2	575	22	AAI18359	Probe #8292 for ge
139	32.4	2.2	575	22	AAI28273	Probe #18206 for g
140	32.4	2.2	575	22	AAI43474	Probe #12160 used
141	32.4	2.2	575	22	AAI57319	Probe #26005 used
142	32.4	2.2	575	22	ABS11588	Human genome-deriv
143	32.4	2.2	575	22	ABS24800	Human genome-deriv
144	32.4	2.2	2465	19	AAH18461	Human granulocytic
145	32.4	2.2	2761	21	AAH61290	Human secreted pro
146	32.4	2.2	5739	24	ABK70271	Human lung cancer
147	32.4	2.2	6125	24	ABK33612	Human immune syste
148	32.4	2.2	6125	24	ABK28277	DNA transcription
149	32.4	2.2	6327	17	AAH32301	Dermatomyositis sp
150	32.4	2.2	6417	24	ABK84696	Human cDNA differe

ALIGNMENTS

RESULT 1

AAZ36101

ID AAZ36101 standard; DNA; 1489 BP.

XX AAZ36101;

XX 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.
 XX Enterohemorrhagic Escherichia coli; EHEC; katp gene; E. coli O157:H7;
 XX IS91; ds.
 XX Escherichia coli.

WO9555908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR01000.

XX 28-APR-1998; 98FR-0005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.
 XX Frechon DTM, Laure FC, Thierry D;
 XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

XX Escherichia coli, particularly serotype O157:H7, used for detecting

XX these bacteria in food

XX Claim 1; Fig 1; 48pp; French.

XX The present sequence is specific to enterohemorrhagic Escherichia coli
 (EHEC). The sequence is 99.9% homologous to the katp gene of E. coli
 O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
 homologous with IS91 of E. coli (nucleotides 1-406 of the present
 sequence). The present sequence is of plasmid origin. Fragments of the
 present sequence are used, as probes and primers, for detection of
 E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
 animal samples, foods or the environment. The fragments are also useful
 for epidemiological studies.

XX SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 other;

Query Match 100.0%; Score 1489; DB 21; Length 1489;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGTCCGGAGATGAAAGCACCACCTGTGTGTACCCATCAGCGTGGTCCGCGAGGCCA 60
 DB 1 CTGCAGTCCGGAGATGAAAGCACCACCTGTGTGTACCCATCAGCGTGGTCCGCGAGGCCA 60

QY 61 TGATTTTGTTCAGACTCAATGACTACCGGAGGACTGAACCTCCGGTGTGTTCTCCA 120
 DB 61 TGATTTTGTTCAGACTCAATGACTACCGGAGGACTGAACCTCCGGTGTGTTCTCCA 120

QY 121 GCCAGTTAAGCCAGCGGTTTCCCTGCTGAAATGTGCGCAAAACGGGGAAGCATCAGAA 180
 DB 121 GCCAGTTAAGCCAGCGGTTTCCCTGCTGAAATGTGCGCAAAACGGGGAAGCATCAGAA 180

QY 181 GGGCGGGGAAGTCCGTCGCGGCGAGTGAACCGTGGCCACTCCGGGAGTACATCCGCC 240
 DB 181 GGGCGGGGAAGTCCGTCGCGGCGAGTGAACCGTGGCCACTCCGGGAGTACATCCGCC 240

QY 241 GCGCGTGATACCGGCAAGAAATGGTCGCAAACTCCCGCTCCGTCGAGCGGCTATTTCAGG 300
 DB 241 GCGCGTGATACCGGCAAGAAATGGTCGCAAACTCCCGCTCCGTCGAGCGGCTATTTCAGG 300

QY 301 ATACCTTGTGTCATCAACACGCTACAAACGAGACAGACAGCTTTTGTGTTCTGACATCCAC 360
 DB 301 ATACCTTGTGTCATCAACACGCTACAAACGAGACAGACAGCTTTTGTGTTCTGACATCCAC 360

301	ATACCCCTTCGTCAACAACAGCTAAACACCAGAACCAAGCTTTTGTTCCTGCATCCAC	360
Qy	AAAGAGGGGAATATTCAGGCTCGCAGCACTCAACGGCACTCGTCAGTTTCGGCTTGGAA	420
Db	AAAGAGGGGAATATTCAGGCTCGCAGCACTCAACGGCACTCGTCAGTTTCGGCTTGGAA	420
Qy	CCCTTAGTATTTTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCCTGTGCATCA	480
Db	CCCTTAGTATTTTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCCTGTGCATCA	480
Qy	ATAAGTTGACTTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTCGTATGA	540
Db	ATAAGTTGACTTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTCGTATGA	540
Qy	TAAAAAAACTCTTCCTGTCTGATCTCTCGCGCTATCGGGAGCTTTTCTACGCTG	600
Db	TAAAAAAACTCTTCCTGTCTGATCTCTCGCGCTATCGGGAGCTTTTCTACGCTG	600
Qy	TAGCCGTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTC	660
Db	TAGCCGTGATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTC	660
Qy	CTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCGTGATTTGATTATGCCACCA	720
Db	CTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCGTGATTTGATTATGCCACCA	720
Qy	GATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTCACAACCT	780
Db	GATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAAAGATTTGCTCACAACCT	780
Qy	CCAGGATTGGTCCCTGCGGATTAATGGTCATTATGGTCCCTTTATTCGTATGGCTT	840
Db	CCAGGATTGGTCCCTGCGGATTAATGGTCATTATGGTCCCTTTATTCGTATGGCTT	840
Qy	GGCAGGTGCGGAAACATACAGACATATGATGGCCGGGAGCGCCAGTGGTGTACG	900
Db	GGCAGGTGCGGAAACATACAGACATATGATGGCCGGGAGCGCCAGTGGTGTACG	900
Qy	AAGCTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAAAGCCGCTCGAT	960
Db	AAGCTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAAAGCCGCTCGAT	960
Qy	TGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCCCTGGGGAGACCTGATGGTCC	1020
Db	TGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCCCTGGGGAGACCTGATGGTCC	1020
Qy	TGACTGGTAAATGTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAA	1080
Db	TGACTGGTAAATGTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAA	1080
Qy	GAGAAGATGACTGGGAGTCGGACCTGTATCTACCTGGGGCTGACAAACAGCCCTCTTGCAG	1140
Db	GAGAAGATGACTGGGAGTCGGACCTGTATCTACCTGGGGCTGACAAACAGCCCTCTTGCAG	1140
Qy	ATAACCGGGATAAAAACGGGAAACTTCAGAAACCTCTTGCCGCCACGACAGATGGGACTTA	1200
Db	ATAACCGGGATAAAAACGGGAAACTTCAGAAACCTCTTGCCGCCACGACAGATGGGACTTA	1200
Qy	TTTTATGTCAATCCTGAAGCCCCGGTGGAAAAACCATCTCTGGCTTCGGCGAAAGATA	1260
Db	TTTTATGTCAATCCTGAAGCCCCGGTGGAAAAACCATCTCTGGCTTCGGCGAAAGATA	1260
Qy	TCAGGGAAGCTTTTTTCACGTATGCCATGGATGAGGAGCTGTGGCCCTCATCGCG	1320
Db	TCAGGGAAGCTTTTTTCACGTATGCCATGGATGAGGAGCTGTGGCCCTCATCGCG	1320
Qy	GAGGGCATACATTTGGTAAAGCATGGTGCAGCGCTCCTGAAAAATGTATTTGGCGCAG	1380
Db	GAGGGCATACATTTGGTAAAGCATGGTGCAGCGCTCCTGAAAAATGTATTTGGCGCAG	1380
Qy	GGCCTGATGGTGCACTGTGGAGGACGAGGACTGGGATGGAAAAATAAATGTGGTACAG	1440
Db	GGCCTGATGGTGCACTGTGGAGGACGAGGACTGGGATGGAAAAATAAATGTGGTACAG	1440

Qy	1441	GAACGGCAAAATATACCATCATCCAGTGGCCCTGGGAAGGAGCCTGGTCGAC	1489
Db	1441	GAACGGCAAAATATACCATCATCCAGTGGCCCTGGGAAGGAGCCTGGTCGAC	1489
RESULT 2			
AAV06555			
ID	AAV06555	standard; DNA; 2238 BP.	
XX	AC	AC	
XX	AAV06555;		
XX	03-JUL-1998	(first entry)	
XX	Microscilla furvens	catalase-53CAL gene.	
XX	Catalase;	epoxidation; hydroxylation; biosensor; paper bleaching	
XX	pasteurisation; ss.		
XX	Microscilla furvens.		
XX	Key	Location/Qualifiers	
XX	CDS	1..2238	
XX	FT	/*tag- a	
XX	FT	/product= "Catalase-53CAL"	
XX	WO9800526-A1.		
XX	08-JAN-1998.		
XX	03-JUL-1997;	97WO-US16513.	
XX	03-JUL-1996;	96US-0674887.	
XX	(RECO-)	RECOMBINANT BIOCATALYSIS INC.	
XX	Adhikary RS,	Robertson DE, Sanyal I;	
XX	WPI; 1998-086953/08.		
XX	P-PSDB; AAW33810.		
XX	New bacterial catalases,	related nucleic acid vectors and	
XX	transformed cells -	used as oxidising agents and for detecting or	
XX	destroying hydrogen peroxide,	e.g. in biosensors	
XX	Claim 3; Fig 2;	35pp; English.	
XX	The present sequence is	of the Microscilla furvens catalase-53	
XX	gene. Fragments of the	gene can be used to identify related seq	
XX	Catalase-53CAL may be	used to catalyse oxidation reactions such	
XX	epoxidation or hydroxylation.	The enzyme can also be used to de	
XX	destroy hydrogen peroxide,	e.g. in connection with glyoxylic ac	
XX	production, biosensors,	contact lens cleaning, pulp/paper bleac	
XX	pasteurisation of dairy	products. Antibodies raised again	
XX	catalase-53CAL can be	used to screen libraries for detection	
XX	purification of cells	containing the enzyme.	
XX	Sequence 2238 BP;	634 A; 545 C; 605 G; 454 T; 0 other;	
XX	Query Match	24.0%; Score 357.6; DB 19; Length 2238;	
XX	Best Local Similarity	66.7%; Pred. No. 2.7e-105;	
XX	Matches 542; Conservative	0; Mismatches 264; Indels 6;	
Qy	683	ATCAAAATCCCTGGGGGCTGATTTTGATATATGCCACCATGATTTCAACAGCTGGAT	
Db	168	ATCGGACCAAAACGACCGGGATTTTGACTATGCGGAAAGATTTTAAAGAAGCTAGAT	
Qy	743	GGCTCTGAAAAAGATATCAAGATTTTCTGCACAACTCCACAGATTTGGTGCCTC	
Db	228	AGCGGGTTAAAAAGGACCTGGCAGCGCTAATGACAGATTCACAGGACTGGTGGCCA	
Qy	803	TTATGTCATATTGTCCTTTCTTTATTTTCGTATGGCTTGGCAGCGTGGCCGGAACAT	

Sequence 2238 BP; 634 A; 545 C; 605 G; 454 T; 0 other;

Query Match	24.0%:	Score	357	6:	DB	19:	Length	2238:
-------------	--------	-------	-----	----	----	-----	--------	-------

Query MatchII

Best Local Similarity 99.7%; Pled: NO: 2./E-105;
Matches 542: Conservative 0: Mismatches 264:

MISSILES 0; MISPLACES 204; INDEXES 6; GAPS 2;

[illegible]

683 ATCAATCCCTGGGGGGCTGATTTTGATTATGCCACCAGATTTC AACAGCTGGATATGA 742

b
168 ATCGGACCCAAACGACCCCGGATTTTGACTATGCGGAGAGACTTTAAGACAGCTACATCTCC 227

108 ATCGGACCCAAACGACCCGGATTTTGTGACTATTGCGGAAGAGTTTAAAGAAGCTAGATCTGGC 227

743 GGCTCTGAAAAAAGATATCAAAGATTGCTGACAACTTCCCAGGATTGGTGGCCCTGCGGA 802

Y
/43 GGCICGAAATAAGATAICAAAGATITGCTGACAACTTCCCAGGATTGGTGCCCTGCGGA 802

b 228 AGCGGTTAAAAGGACCTGGCAGCGCTAATGACAGATTCACAGGACTGGTGGCCAGCAGA 287

2017 07 28 10:00:00

803 TTATGGTCATTATGGTCCTTCTTTATTTCGTATGGCTTGGCACGGTGCCGGAACATACAG 862

[illegible]

896 ATGGCGCCG---GCCCGCGCGATCTGTGTCGGCCCGCCGAAACCCGAGGCTGCTCCGCTGGAGC 952
 1405 AGCAGGACGCGGATGAAATAAATGCTGACAGGAACGCAAAATATACCATCACCA 1464
 953 AGATGGCTTGGCTGGAAGAGCTGTATGCGACCGAACCGTAAGGACCGCATCACCA 1012
 1465 GTGGCCTGGAAGGAGCCTGTGTCGA 1488
 1013 GCGGCATGAGGCTGCTATGGACGA 1036

RESULT 5

AAAX00817
 ID AAAX00817 standard; DNA: 2331 BP.

AC AAAX00817;

DT 26-MAR-1999 (first entry)

DE M. tuberculosis catalase peroxide (katG) gene sequence.

KW Catalase-peroxide; katG; isonicotinic acid hydrazide; tuberculosis;

KW INH; diagnosis; detection; ss.

OS Mycobacterium tuberculosis.

PH Key Location/Qualifiers
 CDS 70..2292

FT /*tag= "katG"

FT /gene= "catalase peroxide"

FT /product= "catalase peroxide"

FT /note= "the start codon is not indicated"

XX WO980585-A1.

XX 12-NOV-1998.

XX 06-MAY-1998; 98WO-US09285.

XX 07-MAY-1997; 97US-0852219.

XX (MAYO-) MAYO FOUNDATION.

XX Cockerill FR, Kline BC, Uhl JR;

XX WPI; 1999-070099/06.

XX P-PSDB; AAW95398.

XX Detection of Mycobacterium tuberculosis - by amplifying katG gene

XX and detecting specific fragment, and optionally identifying

XX INH-resistant strains by detecting specific mutation

XX Claim 3; Fig 7; 83pp; English.

XX The invention relates to a novel method of detecting Mycobacterium
 XX tuberculosis. The method comprises amplifying the DNA in the samples to
 XX generate a detectable amount of amplified DNA comprising a catalase-
 XX peroxide (katG) DNA fragment with sequence of bases 904-1523 of the
 XX M. tuberculosis strain H37Rv MC katG gene sequence. The method optionally
 XX further comprises determining if the katG DNA fragment has a serine to
 XX threonine mutation in codon 315 (S315T mutation), indicative of an
 XX isonicotinic acid hydrazide (INH)-resistant M. tuberculosis strain. The
 XX method can be used to detect M. tuberculosis in biological fluids,
 XX especially human sputum, useful to diagnose tuberculosis. This disease is
 XX a major cause of human morbidity and mortality, and conclusive diagnosis
 XX and subsequent treatment depends on identification of the etiologic agent
 XX M. tuberculosis. INH has been used in tuberculosis treatment, but
 XX INH-resistant strains have emerged; the method allows such drug-resistant
 XX strains to be identified. The present sequence represents the wild-type
 XX M. tuberculosis katG gene sequence.

XX Sequence 2331 BP; 455 A; 734 C; 763 G; 379 T; 0 other;

Query Match 20.0%; Score 298.4; DB 20; Length 2331;
 Best Local Similarity 62.4%; Pred. No. 5.6e-86;
 Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
 QY 688 ATCCCTGGGGGCTGATTTTGGATATGCCACACAGATTTCAACAGCTGGATATGGAGGCTC 747
 DB 236 ACCCGATGGGTGGCGGTTCGACTATGCGCGGAGGTGCGACCATCGACATCTGAGCTT 295
 QY 748 TGAATAAAGATATCAAGATTTTCTGACAACTTCCACAGATTTGGTCCCTGCGGATATG 807
 DB 296 TGACGCGGACATCGAGGAAGTGTACACCATCGCAGCGGTGTGTGCGCCCGGACTAG 355
 QY 808 GTCAATATGCTCTTCTTTATTCGTATGCTGTGGCAGCGTGGCGGAACATACAGACAT 867
 DB 356 GCCACTACGGGCGCTGTTTATCCCGATGGCGTGGCAGCTTCGCGCACCTACCGCATCC 415
 QY 868 ATGATGCGCGGAGCGCGCAGTGGTGTGTCAGCAACGCTTTTGAACGCTGAACAGCTGC 927
 DB 416 ACGACGCGCGCGCGCGCGCGGCGGCGCATGCGAGCGGTTCGCGCGCTTACAGCTGCG 475
 QY 928 CGGATAACCTTAATCTGATAAGCCCGTCTGATTTGCTGTGGCCAGTCAAGAAAAATACG 987
 DB 476 CCGACAACGCCAGCTTGGACAAGGCGCGCGCTGCTGTGGCGGTCAAGAAAGTACG 535
 QY 988 GCTCCAGTATTTCTCTGGGAGACCTGATGCTGCTGACTGCTGATGTTGCTCCCTGATCCA 1047
 DB 536 GCAAGAAGCTCTCATGGCGGACCTGATTTTTCGCGGCAACTGCGCGCTGGAATCGA 595
 QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAGATGACTGGAGTGGAGCTGG 1107
 DB 596 TGGCTTTCAGAGCTTTCGGGTTTCGGCTTCGGCGGTTCGACCATGGGAGCGCGATGAG 655
 QY 1108 TATACTGGGGGCTGACAAAGCCCTTTTCAGATTAACCGGGATA---AAAACGGGAAC 1164
 DB 656 TCTATTGGGCAAGGAAGCACCTGGCTCGGCTGAGCGTTTACAGCGGTAAAGCGGGATC 715
 QY 1165 TTCAGAAACCTCTTTCGGCCACGACAGTGGACTTATTTATGTCATCTCTGAGGCCCCG 1224
 DB 716 TGGAGAACCGCTGGCGCGGTGCGATGAGGCTGATCTACGTGAACCCGAGGGGCCGA 775
 QY 1225 GTGGAACACAGATCTCTGGCTTCGCGGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
 DB 776 ACGGCAACCGGACCCCATGGCGCGGTGCGATTCGCGAGACGTTTCGCGCATGG 835
 QY 1285 CCATGATATGAGGAGACTGTGGCCCTGATCGCGGAGGCGATACATTTGTTAAAGCAC 1344
 DB 836 CCATGAACGACGTCGAACACAGCGCGCTGATCTCGCGGCTCACACTTCGTTAAGACCC 895
 QY 1345 ATGGTCGAGCGCTCTCTGAAAAAATGATTGGCGCAGGCGCTGATGTCGACCTGTGGAGG 1404
 DB 896 ATGGCGCG---GCCCGCGCATCTGGTGGCGCCCGAACCCGAGGCTGCTCCCTGGAGC 952
 QY 1405 AGCAGGACTGGGATGGAATAAATAATGTTGATGAGGAACGCAAAATATACCATCACCA 1464
 DB 953 AGATGGCTTGGCTGGAAGAGCTCGTATGGCACCGGAACCGTAAGGACCGCATCACCA 1012
 QY 1465 GTGGCTGGAAGAGCGCTGTGTCGA 1488
 DB 1013 GCGGCATGAGGCTGCTATGGACGA 1036

RESULT 6

AAI99683/C

ID AAI99683 standard; DNA: 4403765 BP.

XX AAI99683;

XX 15-JAN-2002 (first entry)

XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.

XX Mycobacterium tuberculosis; strain H37Rv; genome;
 XX variation; epidemiology; patient treatment; epidemic monitoring; ds.

XX Mycobacterium tuberculosis.
 XX US6294328-B1.
 XX 25-SEP-2001.
 XX 24-JUN-1998; 98US-0103840.
 XX 24-JUN-1998; 98US-0103840.
 XX (GENO-) INST GENOMIC RES.
 XX Fleischmann RD, White OR, Fraser CM, Venter JC;
 XX WPI; 2001-647261/74.
 XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
 XX determining the nucleotide sequence of the strain at positions in the
 XX genome corresponding to positions where M. tuberculosis strains CDC
 XX 1551 and H37Rv differ
 XX Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
 XX The invention relates to evaluating strain variation within and between
 XX different populations of the tuberculosis bacterial pathogen,
 XX Mycobacterium tuberculosis or related Mycobacterium by determining the
 XX nucleotide sequence of the first strain at positions in the complete
 XX sequence of the genome that correspond to positions that differ in the
 XX nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
 XX H37Rv (AAI99682). The method is useful for evaluating strain variation of
 XX M. tuberculosis and has valuable application in the fields of
 XX M. tuberculosis genetics, epidemiology, patient treatment and epidemic
 XX monitoring.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from USPTO
 XX at seqdata.uspto.gov/sequence.html?docID=6294328B1.
 XX Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
 XX Query Match 20.0%; Score 298.4; DB 22; Length 4403765;
 XX Best Local Similarity 62.4%; Pred. No. 6.2e-84;
 XX Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
 QY 688 ATCCCTGGGGGGCTGATTTTGAATATGCTGACAACTTCCAGGATTGTCCTGCGGATATG 807
 DB 2153203 ACCCGATGGGTGGGGCTTCGACTATGCCCGGAGTTCGGACCATCGACGTTGACGCC 2153144
 QY 748 TGAATAAAGATATCAAGATTTCCTGACAACTTCCAGGATTGTCCTGCGGATATG 807
 DB 2153143 TGACGGGGACATCGAGGAAGTATGATACCACTTCGGACGCTGGTGGCCCGGACTACG 2153084
 QY 808 GTCAATTATGTCCTTTCTTTATTCGATGCTTGGCAGCTTGGCGGAAATACAGGACAT 867
 DB 2153083 GCCACTACGGCGCTGTTTATCCGGATGCGTGGCAGCTGCGGCACTACCGCATCC 2153024
 QY 868 ATGATGCGGGGAGGCGCCAGTGGTGGTCAAGCACTTTTGAACGCTCAAGAGTGGC 927
 DB 2153023 ACACGCGCCGCGCGCGCGGGGGCGGATGACGCGGTTTCGGCGCGCTTAACAGCTGGC 2152964
 QY 928 CGGATACGTTAATCTGGATAAAGCCGCTGATGCTGTCGGCAGTCAAGAAAATACG 987
 DB 2152963 CCGACAAACGCCAGCTTGGACAAGCGCGCGCGCTGCTGTGGCGGTCAAGAAGATPACG 2152904
 QY 988 GCTCCAGTATTTCTGGGGAGACCTGATGCTCTGCTGATTAATTTGCTCTTGAATCCA 1047
 DB 2152903 GCAAGAGCTCTCATGGCGGACCTGATTTTTCGGCGCACTGGCGCTGGAATCGA 2152844
 QY 1048 TGGGATTAAGCGTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTCGAGCTCG 1107
 DB 2152843 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCGGGTCGACCACTGGGAGCCGATGAGG 2152784
 QY 1108 TATACTGGGGGCGCTGACAAACGCTCTTCAGATACCGGGATA---AAACGGGAAAC 1164

DB 2152783 TCATTGGGGCAAGGAAGCCACCTGGCTCGCGGATGAGCTTACAGCGTAAAGCGGATC 2152724
 QY 1165 TTCAGAAACCTCTTGGCGCCACGACATGGGACTTATTTATGTCAATCTCTGAAGGCCCGC 1224
 DB 2152723 TGGAGAACCGCTGGCGCGGCTGCAGATGGGCTGATCTAGTGAACCCGAGGGCGCGA 2152664
 QY 1225 GTGGAACACAGATCTCTCTGGCTTCGCGGAAAGATATCAGGAAAGCTTTTTCAGGTATGG 1284
 DB 2152663 ACGCAACCGGAGCCCATGGCGCGGCTCGACATTCGAGACGCTTTTCGGCGCATGG 2152604
 QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAACAC 1344
 DB 2152603 CCAATGACGACGTCGAAACAGCGCGCTGATCGCGGCTCACACTTTCGGTAAAGACC 2152544
 QY 1345 ATGCTGCAGGCTCTCTCTGAAATAATGATTTGGCGCAGGCGCTGATGTCACCTGTGGAGG 1404
 DB 2152543 ATGGCGCG--GCCGCGCGATCTGTCGCGCCCGAACCGAGGCTGCTCCGCTGGAGC 2152487
 QY 1405 AGCAGGACTGGGATGGAATAATGATGTCGTCAGGAAACGGAATATACCATCACCA 1464
 DB 2152486 AGATGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGACCGGTAAGGACGCGATCACA 2152427
 QY 1465 GTGCGCTGGAAGGAGCTGTGTCGA 1488
 DB 2152426 GCGGCATCGAGGTCGTATGGACGA 2152403
 RESULT 7
 AAI99682/C
 ID AAI99682 standard; DNA; 4411529 BP..
 XX AAI99682;
 XX
 XX 15-JAN-2002 (first entry)
 XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
 XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
 XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
 XX Mycobacterium tuberculosis.
 XX US6294328-B1.
 XX 25-SEP-2001.
 XX 24-JUN-1998; 98US-0103840.
 XX 24-JUN-1998; 98US-0103840.
 XX (GENO-) INST GENOMIC RES.
 XX Fleischmann RD, White OR, Fraser CM, Venter JC;
 XX WPI; 2001-647261/74.
 XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
 XX determining the nucleotide sequence of the strain at positions in the
 XX genome corresponding to positions where M. tuberculosis strains CDC
 XX 1551 and H37Rv differ
 XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
 XX The invention relates to evaluating strain variation within and between
 XX different populations of the tuberculosis bacterial pathogen,
 XX Mycobacterium tuberculosis or related Mycobacterium by determining the
 XX nucleotide sequence of the first strain at positions in the complete
 XX sequence of the genome that correspond to positions that differ in the
 XX nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
 XX H37Rv (AAI99682). The method is useful for evaluating strain variation of
 XX M. tuberculosis and has valuable application in the fields of
 XX M. tuberculosis genetics, epidemiology, patient treatment and epidemic

Db 296 TGACGGGGACATCGAGGAAGTATGACCACTCGCAGCGTGGTGGCCCGCGACTACG 355
 QY 808 GTCAATTATGTCCTTTCTTATTCGATGCGTGGCAGCGTGGCGGAACATACAGACAT 867
 Db 356 GCCACTACGGCGCGCTTTATCCGATGCGTGGCAGCGTGGCGCGCTTACCGCATCC 415
 QY 868 ATGATGGCGGGAGCGCAGTGGTGGTCAAGCGTTTGAACCGCTGAACAGCTGGC 927
 Db 416 ACACCG 475
 QY 928 CGGATAACCTTATCTGGATAAAGCCCGTGGATGCTGGCCAGTCAAGAAAAATACG 987
 Db 476 CGGACAACGCCAGCTTGGACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 535
 QY 988 GCTCCAGTATTCCTGGGAGACCTGATGCTGCTGATGATGATGATGATGATGATGAT 1047
 Db 536 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGCGCACTGCGCGCTGGAATCGA 595
 QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTGGACCTGG 1107
 Db 596 TGGGCTTCAGACGTTGCGGTTGCGGTTGCGCGCGGTGCGACAGTGGGAGCGCGATGAGG 655
 QY 1108 TATACATGGGGCGCTGAACAAGCCCTTTCAGATAACCGGGATA---AAAACGGGAAC 1164
 Db 656 TCTATTGGGCGCAAGGAGCACCTGCTGCGCGATGACGCTTACAGCGGTAAACCGGATC 715
 QY 1165 TTGAGAACTCTTGGCGGCACGACAGATGGGACTTATTTATGTCATCTGAAGGCCCGG 1224
 Db 716 TGGAGAACCCGCTGGCGCGGCTGATGATGGGCTGATCTACGTGAACCCGAGGGCGGA 775
 QY 1225 GTGGAACACCATCTCTGCGTTCCGCGAAGATATCAGGGAAGCTTTTTCAGTATGG 1284
 Db 776 ACGGCAACCGGACCCCATGCGCGCGCGGTGCGACATTCGCGAGAGCTTTGCGCGCATGG 835
 QY 1285 CCATGATGATGAGGAGCTGTGCGCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
 Db 836 CCATGAACGACGTCGAACAGCGCGCTGATCGTGGCGGTACACTTTTCGGTAAGACCC 895
 QY 1345 ATGGTGCAGGCTCTCTGAAATGATTTGGCCGAGGCGCTGATGTCACCTGTGGAGG 1404
 Db 896 ATGGCGCG---GCCGCGCGGATCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 952
 QY 1405 AGCAGGAGCTGGGATGAAATATGATGTTGGTACAGAAACGCAATATACCATCACCA 1464
 Db 953 AGATGGCTGGCTGGAGAGAGCTGATGATGGACCGGACCGGTAAGGACCGCATCACCA 1012
 QY 1465 GTGGCCCTGGAAGGAGCTGGTCTGA 1488
 Db 1013 CCGGCATCGAGGTCGTATGACGA 1036

RESULT 9
 AAT89968
 ID AAT89968 standard; DNA: 2221 BP.

XX AC AAT89968;
 XX AC AAT89968;
 XX AC AAT89968;

20-MAR-1998 (first entry)
 Mycobacterium tuberculosis strain H37Rv variant partial katG gene.
 Tuberculosis; katG gene; isoniazid resistance; INH;
 Isonicotinic acid hydrazide; ss.

XX OS Mycobacterium tuberculosis.

XX PH Key Location/Qualifiers
 XX CDS 1..2217
 XX FT /*tag= a
 XX FT /product= "katG gene"
 XX FT /note= "partial cds"
 XX FT

PN US5688639-A.

XX 18-NOV-1997.
 PD XX
 PF XX
 XX 18-APR-1994; 94US-0228662.
 PR XX
 XX 18-APR-1994; 94US-0228662.
 XX (MAYO-) MAYO FOUNDATION.
 PA XX
 PI Cockerill FR, Kline BC, Uhl JR;
 XX WPI; 1998-007975/01.
 XX
 PT Determination of isoniazid sensitivity of Mycobacterium tuberculosis
 strains - by restriction length polymorphism analysis of katG gene
 PS Disclosure; Column 12-14; 18pp; English.
 CC This partial DNA sequence encodes the katG gene from a variant
 CC strain of Mycobacterium tuberculosis H37Rv which is used in a novel
 CC method to rapidly identify strains of M. tuberculosis which are
 CC resistant to isoniazid (INH, also known as isonicotinic acid hydrazide).
 CC The method involves the use of restriction fragment length polymorphism
 CC (RFLP) analysis to determine if a NciI-MspI restriction site is absent
 CC in the DNA of the strain at the codon corresponding to codon 463 of a
 CC M. tuberculosis katG gene consensus sequence (see AAW31343). The absence
 CC of the site indicates an INH-resistant strain.
 XX SQ
 Sequence 2221 BP; 432 A; 691 C; 735 G; 363 T; 0 other;
 Query Match 18.4%; Score 274.2; DB 19; Length 2221;
 Best Local Similarity 61.4%; Pred. No. 4.3e-78;
 Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;
 QY 688 ATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATATGAGGCTC 747
 Db 176 ACCGATGGTGGCGGCTTCGACTATGCCGGAGGTGCGGACCATCGACTTGCACGCC 235
 QY 748 TGAATAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGTCGCCCTGCGGATATG 807
 Db 236 TGACGGGACATCGAGGAAGTATGATGACCACTCGCAGCGTGGTGGCCCGCGACTACG 295
 QY 808 GTCAATTATGCTCTTTCTTATTCGATGCTGGCAGCGTGGCAGCATACAGGACAT 867
 Db 296 GCCACTACGGCGCGCTTTATCCGATGCGGTGGCAGCTGCGCGCGCATCCGCGATCC 355
 QY 868 ATGATGGCGGGAGCGCGCAGTGGTGGTCAAGCGTTTGAACCGCTGAACAGCTGGC 927
 Db 356 ACACCG 415
 QY 928 CGGATAACCTTATCTGGATAAAGCCCGTGGATGCTGGCCAGTCAAGAAAAATACG 987
 Db 416 CGGACAACGCCAGCTTGGACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
 QY 988 GCTCCAGTATTCCTGGGAGACCTGATGCTGCTGATGATGATGATGATGATGATGAT 1045
 Db 476 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGCGCACTGCGCGCTGGAATC 535
 QY 1046 CATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTGGACCT 1105
 Db 536 GATGGCTTCAGACGTTGCGGTTGCGGTTGCGGTTGCGGTTGCGGTTGCGGTTGCGG 593
 QY 1106 GGTATATGGGGCGCTGACAAAGCCCTTTTCAGATTAACCGGATAAAGCGGAACCT 1165
 Db 594 GGTCTATTGGGCAAGGAGCCACCTGGCTGCGGATGACGCTTACAGCGTAAAGCATCT 653
 QY 1166 TCAGAAACCTCTTGGCGCCACGAGATGGACATTTATTTATGTCATCTCAATCCTGAAGCCCGG 1225
 Db 654 GGAGAACCCGCTGCGCGCGGTGCGATGATGATGATGATGATGATGATGATGATGAT 713
 QY 1226 TGGAAAAACCATGCTCTGCTGCTGCGGAAAGATATCAGGGAAGCTTTTTCAGTATGGC 1285
 Db 714 CGGCAACCCGACCCCATGGCGCGGTGCGACATTCGCGGAGACGTTTTCGCGCGCATGGC 773

XX DT 11-MAY-1999 (first entry)
 XX DE M.tuberculosis katG gene.
 XX DE katG; catalase; peroxidase; enzyme; isonicotinic acid hydrazide;
 KW K isoniazid; INH; diagnosis; resistance; ss.
 XX OS Mycobacterium tuberculosis.
 XX FH Key Location/Qualifiers
 FT CDS 1979..4186
 FT /*tag= a
 FT /gene= "katG"
 FT /product= "catalase-peroxidase"
 XX PN US5871912-A.
 XX PD 16-FEB-1999.
 XX PF 02-JUN-1995; 95US-0459499.
 XX PR 02-JUN-1995; 95US-0459499.
 PR 30-APR-1992; 92US-0875940.
 PR 14-AUG-1992; 92US-0929206.
 PR 11-MAR-1993; 93US-0029655.
 XX (INSP) INST PASTEUR.
 XX Cole ST, Heym B, Young DB, Zhang Y;
 WPI; 1999-166622/14.
 DR P-PSDB; ANW78363.
 XX New nucleic acid encoding catalase-peroxidase HPI - useful for rapid
 PT identification of isoniazid resistant Mycobacterium tuberculosis
 XX Disclosure; Fig 10A; 46pp; English.
 XX This sequence represents the Mycobacterium tuberculosis katG gene, which
 CC encodes a catalase-peroxidase enzyme isolated from a M. tuberculosis
 CC strain resistant to isonicotinic acid hydrazide (isoniazid or INH).
 CC The sequence is useful in the diagnosis of tuberculosis caused by
 CC INH-resistant M. tuberculosis and in the identification of other
 CC INH-resistant strains.
 XX Sequence 4795 BP; 892 A; 1532 C; 1559 G; 812 T; 0 other;
 SQ
 Query Match 18.4%; Score 274.2; DB 20; Length 4795;
 Best Local Similarity 61.4%; Pred. No. 6.9e-78;
 Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;
 QY 688 ATCCCTGGGGGCGTATTGATATGCCACAGATTTCACAGCTGGATATGAGGCTC 747
 DB 2145 ACCCGATGGTGGCGGCTTCGACTATGCCGCGAGGTCCGACCGCTCGACTTACGCC 2204
 QY 748 TGAATAAAGATATCAAGATTCTGCTGACAACTTCCAGGATTTGGTCCCTCGCGGATATG 807
 DB 2205 TGACGGGGACATCGAGGAGTATGATGACCACTCGCAGCGTGGTGGCGCGACTACG 2264
 QY 808 GTCAATATGCTCTTCTTTATTCGTATGCTGGCAGCGTCCGGAACATACAGGACAT 867
 DB 2265 GCCACTACGGCGCTGTTTATCCGGATGGCGTGGCAGCTCGCGACCTCCCGACTCC 2324
 QY 868 ATGATGCCCGGGAGGCGCGAGTGGTGTGAGCAAGCTTTTGAACGCTGAACAGCTGGC 927
 DB 2325 ACCAGCGCGCGGGCGCGCGGGCGGATGATGACGGTTCGCGCGCTTAACAGCTGGC 2384
 QY 928 CGGATACGTTAATCTGGATTAAGCCCGCTGATGCTGGCGCAGTCAAGAAAATACG 987
 DB 2385 CCGACACGCGAGCTTGGACAAGCGCGCGCGTGTGTGGCGGTCAAGAAGATACG 2444
 QY 988 GCTCCAGTATTTCTGGGGAGACCTGATGCTCTGACTGGTAA--TGTTCCCTTGNATC 1045

DB 2445 GCAGAAGCTCTCATGGGGGACCTGATTTGTCGGCCCAACCGTGGCTCGGAATC 2504
 QY 1046 CATGGGATTAAACGCTGGGATTGCTGGCGAAGAGAGATGACTGGGAGTGGACCT 1105
 DB 2505 GATGGGCTTCAAGACGTTCCGGTTTCGGTTCGG--GCGTCGACACAGTGGGAGACCGATGA 2562
 QY 1106 GGTATCTACTGGGGCTTGACAAAGCTCTTTGAGATACCGGGATAAACCGGAAACT 1165
 DB 2563 GGTCTATTGGGGCAAGGAAGCCACCTGGCTCGCGGATGACGGTTACAGCGTAAGCGTCT 2622
 QY 1166 TCAGAAACCTCTTTCGCGCCACGACAGATGGGACTTATTATGTCATCTCAAGGCCCGG 1225
 DB 2623 GGAAACCCGCTGGCGCGGTGAGATGGGCTGATCTAGCTGAACCCGGAGCGCGGAA 2682
 QY 1226 TGGAAACACAGATCCTCTGGCTTCGCGAAAGATATCAGGGAAGCTTTTTCAGGTATGGC 1285
 DB 2683 CGCAACCCGACCCCATGGCGCGGTCGACATTCGGGAGACGTTTCGGCGCATGGC 2742
 QY 1286 CATGGATGATGAGGAGACTGTGCCCTGATCGCGGAGGAGCATATTTGGTAAAGACA 1345
 DB 2743 CATGAACGAGCTGAAACACGCGGCTGATCGTCGCGGTCACACTTTTCGGTAAGACCA 2802
 QY 1346 TGGTGCAGCGCTCTCTGAAATGATTTGGCGCAGGCGCTGATGTCACCTCTGGAGGA 1405
 DB 2803 TGGCGCGG---GCCCGCGGATCTGGTCCGCCCGCAACCGGAGGCTGCTCCGCTGGAGCA 2859
 QY 1406 GCAGGAGCTGGGATGGAATAATAATGTTGTCAGGAAACGCAATATACCATCACCAG 1465
 DB 2860 GATGGCTTGGGCTGGAAGAGCTGATGTCGACCGCAACCGGTAAAGACCGCATCACCAG 2919
 QY 1466 TGGCTGGGAAGAGCGCTGGTCGA 1488
 DB 2920 CGGCATCGAGGTCGTATGGACGA 2942
 RESULT 12
 AAQ06815
 ID AAQ06815-standard; DNA; 2196 BP.
 AC AAQ06815;
 DT 06-MAR-1991 (first entry)
 XX Sequence encoding heat resistant peroxidase.
 DE pod10; E.coli UM228; ds.
 KW Bacillus stearothermophilus.
 OS
 XX Key Location/Qualifiers
 FT CDS 1..2193
 FT /*tag= a
 XX JP02268684-A.
 XX 02-NOV-1990.
 XX 07-APR-1989; 89JP-0089469.
 XX 07-APR-1989; 89JP-0089469.
 XX (TOYM) TOYOBO KK.
 XX WPI; 1990-372008/50.
 DR P-PSDB; AAR08205.
 XX DNA contg. genetic information of heat resistant peroxidase -
 PT prepd. by culturing transformant and collecting heat resistant
 PT peroxidase
 XX Disclosure; Fig 2; 8pp; Japanese.
 PS
 XX

CC Sequence may be used to construct plasmid p010 used to transform
 CC E.Coli UM228. The transformed expression system may be used to
 CC produce heat resistant peroxidase in large quantities.

XX
 SQ Sequence 2196 BP; 576 A; 564 C; 633 G; 423 T; 0 other;

Query Match 16.8%; Score 250.4; DB 11; Length 2196;
 Best Local Similarity 59.0%; Pred. No. 2.5e-70;
 Matches 524; Conservative 0; Mismatches 351; Indels 13; Gaps 5;

QY	607	CTGATAAAGAGACTCAAAATTTCTACTATCCAGAAACACCTGGATTAACTCCCTCTGA	666
DB	65	CTTCAATCGAACGACGAACAGACTGGTGGCGCAACACAGCTGAACCTTAAGCAATTC	124
QY	667	GATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGTATTGATGACCAAGATTC	726
DB	125	ATCAACATGACGAAACAGNATCTCATGATGAGAGTTCACTATGCTGAGAGTTTC	184
QY	727	AACAGCTGGATGATGAGGCTCTGAAAGAAAGATATCAAAAGATTTGCTGCAACATTC	786
DB	185	AAAAACTAGACTATTGGGCGCTCAAGAAGATTTGCGCAAACTGATGACGGAAGCAAG	244
QY	787	ATTGGTGCCTCGGATATGATGATGATGATGATGATGATGATGATGATGATGATG	846
DB	245	ACTGGTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG	304
QY	847	GTGCGGAACATACAGGACATATGATGCGCGGGGAGCGCAGTGGTGCAGCAAGTT	906
DB	305	CAGCTGGCAGCTACCGCATCGCCAGCGCGCGCGCGCTTCGACCGGACGACGGCT	364
QY	907	TTGAACCGCTGAACAGCTGGCGGATAACTGTTAATCTGATGATAAGCCGCTGATG	966
DB	365	TTGGCGGCTTAACAGCTGGCGGACCAACGCACTTTGATGATAAGCGCGGCTTGT	423
QY	967	GSCAGTCAAGAAATACGCTCCAGTATTTCTGGGAGACCTGATGCTCTGCTACTG	1026
DB	424	GGCGGATCAAAAGAAATACGCGAACAATCTCTGGCGGATTTGTTTCAATTTGGCG	483
QY	1027	GTAATGTTGCCCTTGAATCAATGAGGATTTAAAGCGCTGGGATTTGCTGGCGGAAGAAG	1086
DB	484	GCAATGTCGCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATG	543
QY	1087	ATGACTGGGAGTGC---GACCTGATATGATGATGATGATGATGATGATGATGATG	1143
DB	544	ACGCTGGCATCCGGAAGAGACGCTTTATTGGGATCGGAAAGAGTGGCTCGCCCTG	603
QY	1144	ACCGGATAAAGCG---GGAACCTTCAGAAACCTCTTGGCGGACGAGATGGACTAT	1201
DB	604	AACGCTATTCGGGTATCGCGACCTCGAAACCCGCTCGCGCGCTGCAATGCGGTAT	663
QY	1202	TTATGCTAATCCTGAAGCCCGGTTGGAACACCATGCTCTGCTGCTTCCGGAAGATAT	1261
DB	664	CTAGCTCAACCCAGAGGCGGACGCGACGCGGATCCAAAGACGAGGGGATATC---	719
QY	1262	CAGGAACCTTTTCACTATGCGCATGATGATGATGATGATGATGATGATGATGATG	1321
DB	720	CGCAGAGAGCTTCGCGCATGATGATGATGATGATGATGATGATGATGATGATG	779
QY	1322	AGGGCATACATTTGTTAAGACACATGTTGCGAGCTCTCTCGAAATATGTTTGGCGCAG	1381
DB	780	CGGTCTATGCTGGAAAGCGCACCGCGGCG---GCCCTGCCAGCGCTGCTCCCGA	836
QY	1382	GCCTGATGCTGACCTGTGGAGGACGAGGACTGGGATGGAATAAATATGTTGATACAG	1441
DB	837	GCCGGAAGCCCGCGGATTAAGCGCAAGGCTGGGATGATGATGATGATGATGATG	896
QY	1442	AAACGGCAATATACCATCAGTGGCTCGAAGGAGCCCTGGTCGAC	1489
DB	897	GAAAGGAGCGATACGATCAAGCGGATTAAGCGGCTTTGAGCGCTTTGAGCGCC	944

RESULT 13
 AAV06554

ID AAV06554 standard; DNA; 2262 BP.
 AC AAV06554;
 XX
 DT 03-JUL-1998 (first entry)
 XX
 DE Alcaligenes (Deleya) aquamarinus catalase-64CA2 gene.
 XX
 KW Catalase; epoxidation; hydroxylation; biosensor; paper bleaching;
 KW pasteurisation; ss.
 XX
 OS Alcaligenes aquamarinus.
 XX
 FH Key Location/Qualifiers
 FT 1..2262
 FT /*tag- a
 FT /product= "Catalase-64CA2"

W09800526-A1.

08-JAN-1998.

03-JUL-1997; 97WO-US16513.

03-JUL-1996; 96US-0674887.

(RECO-) RECOMBINANT BIOCATALYSIS INC.

Adhikary RS, Robertson DE, Sanyal I;

WPI; 1998-086953/08.

P-PSDB; AAW33809.

New bacterial catalases, related nucleic acid vectors and transformed cells - used as oxidising agents and for detecting or destroying hydrogen peroxide, e.g. in biosensors

Claim 3; Fig 1; 35pp; English.

The present sequence is of the Alcaligenes aquamarinus catalase-64CA2 gene. Fragments of the gene can be used to identify related sequences. Catalase-64CA2 may be used to catalyse oxidation reactions such as epoxidation or hydroxylation. The enzyme can also be used to detect or destroy hydrogen peroxide, e.g. in connection with glyoxylic acid production, biosensors, contact lens cleaning, pulp/paper bleaching and pasteurisation of dairy products. Antibodies raised against catalase-64CA2 can be used to screen libraries for detection and purification of cells containing the enzyme.

Sequence 2262 BP; 553 A; 612 C; 663 G; 434 T; 0 other;

Query Match 15.1%; Score 224.6; DB 19; Length 2262;
 Best Local Similarity 57.4%; Pred. No. 6.5e-62;
 Matches 476; Conservative 0; Mismatches 329; Indels 24; Gaps 3;

QY	682	AATCAATCCCTCGGGGCTGATTTGATATGCCACAGATTTCAACAGCTGGATATGG	741
DB	218	AATCAGACCCGATGGATCCGGATTTCAACTACCGTGAAGAAGTACGCAAGCTGATTCG	277
QY	742	AGCTCTGAAAAGATATCAAGATTTCTGACAACTTCCAGGATTTGGTCCCTGCGG	801
DB	278	ACGCGCTGAAGAAGATGTCACGCGTTGATGACCATAGCAAGTGTGGCCCGCTG	337
QY	802	ATTATGTCATTATGTCCTTTCTTTATTCGATGCTTGGCAGCGTGGCAACATACA	861
DB	338	ACTGGGGCACTACGGCGGTTTGATGATCGTATGCTTGGCACTCCGCTGGCACTACC	397
QY	862	GGACATATGATGCGGGGAGCGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	921
DB	398	GTATTGCTGATGCGCGTGGGGGCGGTGACCGCAAGAGCGCGCTTTGCACCGCTCACT	457
QY	922	GCTGGCGGATACGTTAATCTCGATAAAGCCGCTGATTTGCTGCGCAGTCAAGAAA	981

Db 458 CCTGGCGGACACCTCAGCTGGATTAAGCGCGCGCTGCTGCTGGCCGATCAAGA 517
 QY 982 AATACGGCTCAGATATTCTGGGAGACCTGATGCTCTGACTGTAATGTTGCCCTTG 1041
 Db 518 AGTACGGCAACAAATCAGCTGGCAGACCTGATGATTTGCTGGCAGCGTGGCTTATG 577
 QY 1042 AATCCATGGGATTTAAACGCTGGGATTTGCTGGCGAAGAGATGACTGGAGTCGG 1101
 Db 578 AGTCCATGGCTTACCTGCTTACGGCTTCTTTTCGGCGCGCTGCGATATTTGGGAACCG 637
 QY 1102 AACTGG---TATACTGGGGCCTTGACAACAAGCCTCTTGACAGAT----- 1142
 Db 638 AAAAGATATCTACTGGGTGACGAAAAGAGTGGCTGGCACCTTCTGACAACTGACG 697
 QY 1143 --AACCGGATTAACCGGAACTTCAGAAACCTCTTGGCGCCGACGAGATGGGACTTA 1200
 Db 698 GCGACGTGAACAAGCCAGAGACCATGGAACACCGCTGGCGGTGTCCAAATGGGTCTGA 757
 QY 1201 TTTATGTCATCTGAAGCGCCGCTGGAAACACAGATCTCTGCTTCCGCGAAAGATA 1260
 Db 758 TCTATGTGAACCGGAGGTGTTAAGCGCCACCTGATCGCTGAGAACCGCACAGG 817
 QY 1261 TCAGGAAGCTTTTTCACGTATGGCCATGATGATGAGGAGACTGTGGCCTGATCGCGG 1320
 Db 818 TACTTGAACCTTTCGCGCTATGGCGATGAACGACGAAACACCGCAGCCCTCACAGCTG 877
 QY 1321 GAGGCGATACATTTGTTAAGCACATGTTGACGCTCTCTGTAATAATGTATTGGCGCAG 1380
 Db 878 GCGGCCACACCGCTGCGGTATGTCACGGTAATGGCAATGCTCTGCTGCTGCTG 934
 QY 1381 GGCTGATGTGACCTGTGGAGGACGAGGACTGGGATGGAAATAATGTGGTACAG 1440
 Db 935 ACCAAAGCTCTGACGTGTAACACGAGGCTTAGGTGGGCAACCCCAACATCGAGG 994
 QY 1441 GAAAGCGCAATATACCATCACAGTGGCTTGAAGGAGCGCTGGTGAC 1489
 Db 995 GCAAGCAAGCAACGCGTGACCTCGGCTATCGAAGGTGCTTGGACCAC 1043

RESULT 14

AA49864
 ID AAA49864 standard; DNA; 660 BP.
 AC AAA49864;
 XX
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE Mycobacterium tuberculosis katG gene (isoniazid resistance).
 XX
 KW Antibiotic resistance; katG gene; isoniazid resistance; ss.
 XX
 OS Mycobacterium tuberculosis.
 XX
 FH Key Location/Qualifiers
 FT primer_bind complement(41..60)
 FT /tag- a
 FT /note- "primer of AAA49827"
 FT 231..250
 FT /tag- b
 FT /note- "primer of AAA49828"
 XX
 XX WO200036142-A1.
 XX
 XX 22-JUN-2000.
 PD
 XX
 XX 10-DEC-1999; 99WO-CA01177.
 XX
 XX 11-DEC-1998; 98US-0111794.
 PR
 XX (VISI-) VISIBLE GENETICS INC.
 PA
 XX Shipman R;
 PI
 XX

DR XX WPI; 2000-431611/37.
 PT Method for the detection and characterization of Mycobacterium
 XX tuberculosis with antibiotic resistance in a sample -
 PS Disclosure; Page 5-6; 43pp; English.
 CC The present sequence is that of the Mycobacterium tuberculosis
 CC katG (isoniazid resistance) gene (bp661-1320). Amplification and
 CC cycle sequencing primers (see AAA49823-62) are used for the detection
 CC and analysis of antibiotic resistance-associated mutations in
 CC defined regions of rpoB (rifampin), katG (isoniazid), oxyR-aphC PR
 CC (isoniazid), mabA (isoniazid), rpsL/s12 (streptomycin), 16S/rrs
 CC (streptomycin), emmB (ethambutol), pncA (pyrazinamide), gyrA
 CC (ciprofloxacin) and 23S (azithromycin) genes of M. tuberculosis.
 CC These primers can be used in a method for the detection and
 CC characterization of M. tuberculosis present in a sputum sample.
 CC The method involves performing a sequencing procedure, with or
 CC without prior amplification, to detect the presence of M.
 CC tuberculosis, and if present to evaluate the rpoB, katG, rpsL/s12
 CC and 23S genes for the presence of antibiotic-inducing mutations.
 CC If M. tuberculosis is detected, a second sequencing procedure is
 CC performed on the sample to evaluate additional genes for the
 CC presence of antibiotic resistance-inducing mutations. Genotypic
 CC tests are rapid, sensitive and accurate providing information as to
 CC antibiotic treatment options.
 XX
 SQ Sequence 660 BP; 127 A; 206 C; 227 G; 100 T; 0 other;
 Query Match 8-2%; Score 122.8; DB 21; Length 660;
 Best Local Similarity 62.6%; Pred. No. 4.9e-29;
 Matches 209; Conservative 0; Mismatches 122; Indels 3; Gaps 1;
 QY 1155 AAGCGGAACCTTCAGAAACCTCTTGGCCGACGATGGGACTTATTTATGTCAATCCT 1214
 Db 26 AAGCGGATCTGGAGAACCGCTGGCGGTCGATGGGCTGATCTACGTGAACCCG 85
 QY 1215 GAAGGCCCGGTGAAACACGATCTCTGGCTTCCGCGAAGATATCAGGGAAGCTTTT 1274
 Db 86 GAGGGGCGAAGAACCGGACCGGACCCATGGCGCGGCTGCACATTCGCGAGACGTTT 145
 QY 1275 TCACGTATGGCCATGATGATGAGGAGACTGTGGCTGATCGCGGAGGCGCATACATTT 1334
 Db 146 CGCGCGATGGCCATGATGACGACGTCGAAACGCGGCTGATCTGCGCGCTACACTTC 205
 QY 1335 GGTAAAGCACATGGTGCAGCGCTCTCTGAAATAATGTATGGCGCAGGCGCTGATGTGCA 1394
 Db 206 GGTAAAGCACATGGCGCGG---CCCGCGCGATCTGTGCGCCCGAACCCGAGGCTGCT 262
 QY 1395 CCTGTGAGGAGCAGGAGCTGGGATGGAATAATGTTGTTACAGGAACCGCAATAT 1454
 Db 263 CGCTGGAGCAGATGGGCTTGGCTGGAGAGCTCGTATGGCACCGGACCGGTAAAGAC 322
 QY 1455 ACCATCACCAGTGGCCTGGAAAGAGCGCTGGTCCA 1488
 Db 323 GCGATCACCAGCGGCTGAGGTCGTATGGACGA 356
 RESULT 15
 AA497490
 ID AA497490 standard; cDNA; 1852 BP.
 AC AA497490;
 XX
 XX 13-MAR-2001 (first entry)
 DT
 XX
 XX Fusarium venenatum EST SEQ ID NO:13.
 DE
 XX Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; Identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.

PN EP1076099-A2.
 PD 14-FEB-2001.
 PF 02-AUG-2000; 2000EP-0306563.
 PR 03-AUG-1999; 99JP-0220357.
 XX (NISN) NISSHINBO IND INC.
 PA (SYST-) SYSTEM RES INC.
 XX Suzuki Y, Nishida M, Takenishi S;
 PI WPI; 2001-246696/26.
 DR
 XX
 PT New oligonucleotides, nucleic acid probes and primers are useful for
 PT differentiating drug-resistance and determining infection with tubercle
 PT bacilli -
 XX
 PS Disclosure; Page 61-62; 114pp; English.
 CC
 CC The present invention relates to oligonucleotides based on nucleotide
 CC sequences obtained from both wild-type tubercle bacilli (wTB) that are
 CC susceptible to a drug and mutant-type tubercle bacilli (mTB) that are
 CC resistant to a drug. The drugs used in the present invention are
 CC rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
 CC ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
 CC rrs gene is responsible for resistance to SM and KM; the rpsL gene is
 CC responsible for resistance to SM; the inhA gene is responsible for
 CC resistance to INH; the katG gene is responsible for resistance to INH;
 CC and the embB gene is responsible for resistance to EB. The present
 CC invention also relates to nucleic acid probes having part of a nucleotide
 CC sequence of tubercle bacilli (TB) responsible for drug resistance and
 CC primers used to generate the probes. The present sequence is an
 CC oligonucleotide of the present invention. The oligonucleotides of the
 CC present invention can be used to enable the differentiation of drug
 CC resistance and the determination of infection with tubercle bacilli
 CC simultaneously.
 XX
 SQ Sequence 150 BP; 29 A; 51 C; 53 G; 17 T; 0 other;

Query Match 5.1%; Score 75.6; DB 22; Length 150;
 Best Local Similarity 71.7%; Pred. NO. 4.9e-14;
 Matches 99; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 852 GGAACATACAGACATATGATGCGCGGGAGCGCCAGTGTGTGTCAGCAACGTTTGA 911
 DB 1 GGCACCTACCGATCCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 QY 912 CGCTGAACAGCTGCCGGATACGTTAATCTGGATAAGCCGCTGATGCTGTGCCCA 971
 DB 61 CGGCTTACAGCTGGCCGACAGCCAGCTTGGACAAGGCCCGCGCGCGCGCGCGCG 120
 QY 972 GTCAGAAAATAACGCG 989
 DB 121 GTCAGAGAAGATACGCG 138

RESULT 19
 AAF95161
 ID AAF95161 standard; DNA; 150 BP.
 XX
 AC AAF95161;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE katG gene probe #3.
 XX
 KW Tubercle bacillus; drug sensitivity; drug resistance; rifampicin;
 KW streptomycin; kanamycin; isoniazid; ethambutol; rpoB gene; rrs gene;
 KW rpsL gene; inhA gene; katG gene; embB gene; probe; PCR primer; ss.
 XX
 OS Mycobacterium tuberculosis.

XX EP1076099-A2.
 XX 14-FEB-2001.
 XX 02-AUG-2000; 2000EP-0306563.
 XX 03-AUG-1999; 99JP-0220357.
 XX (NISN) NISSHINBO IND INC.
 XX (SYST-) SYSTEM RES INC.
 XX Suzuki Y, Nishida M, Takenishi S;
 XX WPI; 2001-246696/26.
 XX
 XX New oligonucleotides, nucleic acid probes and primers are useful for
 XX differentiating drug-resistance and determining infection with tubercle
 XX bacilli -
 XX
 PS Disclosure; Page 62; 114pp; English.
 CC
 CC The present invention relates to oligonucleotides based on nucleotide
 CC sequences obtained from both wild-type tubercle bacilli (wTB) that are
 CC susceptible to a drug and mutant-type tubercle bacilli (mTB) that are
 CC resistant to a drug. The drugs used in the present invention are
 CC rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
 CC ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
 CC rrs gene is responsible for resistance to SM and KM; the rpsL gene is
 CC responsible for resistance to SM; the inhA gene is responsible for
 CC resistance to INH; the katG gene is responsible for resistance to INH;
 CC and the embB gene is responsible for resistance to EB. The present
 CC invention also relates to nucleic acid probes having part of a nucleotide
 CC sequence of tubercle bacilli (TB) responsible for drug resistance and
 CC primers used to generate the probes. The present sequence is an
 CC oligonucleotide of the present invention. The oligonucleotides of the
 CC present invention can be used to enable the differentiation of drug
 CC resistance and the determination of infection with tubercle bacilli
 CC simultaneously.
 XX
 SQ Sequence 150 BP; 28 A; 37 C; 54 G; 31 T; 0 other;
 Query Match 4.5%; Score 66.8; DB 22; Length 150;
 Best Local Similarity 65.3%; Pred. NO. 3.6e-11;
 Matches 98; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 QY 962 GCTGTGCCAGTCAAGAAAATACGCTCCAGTATTTCTCTGGGAGACCTGATGCTCT 1021
 DB 1 GCTGTGCCGCTCAAGAAAGATACGCAAGAGCTCTCATGGCGGACCTGATTTT 60
 QY 1022 GACTGGTAATGTGCGCTTGAATCCATGGATTAAACGCTGGGATTTCTGGCGGAAG 1081
 DB 61 CGCGGCAACTGCGCGCTGGAATCGATGGCTTCAAGACGCTTGGGCTTCGGCGG 120
 QY 1082 AGAAGATGACTGGGAGTCGACCTGCTATA 1111
 DB 121 GGTCGACAGTGGGAGCCGCGATGAGTCTA 150

RESULT 20
 AAF95159
 ID AAF95159 standard; DNA; 150 BP.
 XX
 AC AAF95159;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE katG gene probe #1.

XX Tubercle bacillus; drug sensitivity; drug resistance; rifampicin;
 KW streptomycin; kanamycin; isoniazid; ethambutol; rpoB gene; rrs gene;
 KW rpsL gene; inhA gene; katG gene; embB gene; probe; PCR primer; ss.
 XX

OS Mycobacterium tuberculosis.
 PN EP1076099-A2.
 XX
 XX 14-FEB-2001.
 XX
 XX 02-AUG-2000; 2000EP-0306563.
 XX
 XX 03-AUG-1999; 99JP-0220357.
 XX
 XX (NISN) NISSHINBO IND INC.
 PA (SYST-) SYSTEM RES INC.
 PA
 XX Suzuki Y, Nishida M, Takenishi S;
 PI
 XX WPI; 2001-246696/26.
 DR
 XX
 XX New oligonucleotides, nucleic acid probes and primers are useful for
 PT differentiating drug-resistance and determining infection with tubercle
 PT bacilli -
 PT
 XX Disclosure; Page 61; 114pp; English.
 PS
 XX The present invention relates to oligonucleotides based on nucleotide
 CC sequences obtained from both wild-type tubercle bacilli (wtTB) that are
 CC susceptible to a drug and mutant-type tubercle bacilli (mtTB) that are
 CC resistant to a drug. The drugs used in the present invention are
 CC rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
 CC ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
 CC rrs gene is responsible for resistance to SM and KM; the rpsL gene is
 CC responsible for resistance to SM; the inhA gene is responsible for
 CC resistance to INH; the katG gene is responsible for resistance to INH;
 CC and the embB gene is responsible for resistance to EB. The present
 CC invention also relates to nucleic acid probes having part of a nucleotide
 CC sequence of tubercle bacilli (TB) responsible for drug resistance and
 CC primers used to generate the probes. The present sequence is an
 CC oligonucleotide of the present invention. The oligonucleotides of the
 CC present invention can be used to enable the differentiation of drug
 CC resistance and the determination of infection with tubercle bacilli
 CC simultaneously.
 XX
 SQ Sequence 150 BP; 26 A; 53 C; 48 G; 23 T; 0 other;
 Query Match 4.3%; Score 63.8; DB 22; Length 150;
 Best Local Similarity 66.2%; Pred. No. 3.4e-10;
 Matches 92; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 733 TGGATATGGAGGCTCTGAAAAAGATATCAAGATTTCGTGCAACTTCCAGGATTGGT 792
 DB 12 TCGACGTTGACGCCCTGACCGGGACATCGAGAGTATGATGACCTCGCAGCGTGGT 71
 QY 793 GCCCTGCGGATTATGTCATATATGTCCTTTTATTCGTTATGCTTGGTGGCAGCGTCCG 852
 DB 72 GCGCCGCGGACTACGGCCACTACGGCGCGCTGTTTATCCGGATGGCGTGGCGTCCG 131
 QY 853 GAACATACAGGACATATGA 871
 DB 132 GCACCTACCGCATCCACGA 150
 RESULT 21
 ABQ23236/c
 ID ABQ23236 standard; DNA; 538 BP.
 XX
 AC ABQ23236;
 XX
 XX 12-JUL-2002 (first entry)
 DE
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 9827.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW

KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 XX 07-MAR-2002.
 PD
 XX
 XX 01-SEP-2001; 2001WO-EP10074.
 PF
 XX
 XX 01-SEP-2000; 2000DE-1043826.
 PR
 XX 05-SEP-2000; 2000DE-1044543.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;
 PI
 XX WPI; 2002-371829/40.
 DR
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 PT
 XX Claim 12; 56pp + Sequence Listing; 56pp; German.
 PS
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 538 BP; 111 A; 47 C; 142 G; 238 T; 0 other;
 Query Match 3.2%; Score 47.4; DB 24; Length 538;
 Best Local Similarity 45.9%; Pred. No. 0.00017;
 Matches 162; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
 QY 1014 ATGCTCTGACTGTGTAATGTGCGCTTGAATCCATGGATTAAACGCTGGGATTGCT 1073
 DB 524 ATTATCTCGGATACGTAAACGCTAAACAACTCCGACTTCGTACCTTCGATTTAAT 465
 QY 1074 GCGGGAAGAAGATGACTGGAGTGGACCTGGTATGATGCGGCGCTGACAAACAGCCT 1133
 DB 464 ACCGATCGTAAAAACGCTCTAAAAACCGAATCTAAATATTAACTAAATAATAAAAAACC 405
 QY 1134 CTTGCAGATAACCGGGATAAAACGGGAACCTTCAGAAACCTCTTGGCCGCCACGACATG 1193
 DB 404 TAACTAACTACCGCTCATCCGNAACGCTAACGAAACACCGCTTAATATCAACCGAATA 345
 QY 1194 GGACTTATTATGTCATCTGAAGCCCGGGTGAACACACAGATCTCTGGCTTCGCG 1253
 DB 344 AATCTAATTTACGTTAACCGGAAAAACCCGAATCAACACGACGACCGCTTCTACGACA 285
 QY 1254 AAGATATACGGAAGCTTTTTCACGTTATGCCATGGATGATGAGGATGTGGCCCTG 1313
 DB 284 ACAATATCCGCGCGACCTTTCGACACATATAACATAAACGACGAAAAACCGTAACGCTA 225
 QY 1314 ATCGCGGAGGCGCATACATTTGGTAAAGACATGGTCCACCGCTCTCTGAAAA 1366

CC	resistance and the determination of infection with tubercle bacilli simultaneously.
CC	
XX	
SO	Sequence 150 BP; 26 A; 49 C; 51 G; 24 T; 0 other;
	Query Match 3.18; Score 46.4; DB 22; Length 150;
	Best Local Similarity 61.2%; Pred. No. 0.00016;
	Matches 93; Conservative 0; Mismatches 56; Indels 3; Gaps 1;
QY	1260 ATCAGGGAACCTTTTTCACGTATGGCCATCGATGATGAGGACTGTGCCCTGATCGCG 1319
DB	2 ATTCCGAGAGCTTTCGGCGCATGGCCATTGAACGACGTCAAAACAGCGCGCTGATGCTC 61
QY	1320 GGAGGGCAATACATTTGGTAAGAACAATGTGGTGACGGTCTCCTGAAAATGTATTGGCGCA 1379
DB	62 GGGGGTTCACACTTTCGTGTAGACCATTGGCGCGGGCCCCGCATCTGG---TCGGCCCC 118
QY	1380 GGGCCTGATGGTGCACTGTGGAGGACGAGG 1411
DB	119 GAACCCGAGGCTGCTCCGCTGGAGCAGATGG 150

RESULT 24	
AAFl1909	
ID	AAFl1909 standard; CDNA; 1358 Bp.
XX	
XX	
AC	AAFl1909;
XX	
DT	13-MAR-2001 (first entry)
XX	
DE	Aspergillus oryzae EST SEQ ID NO:4432.
XX	
XX	
KW	Multiple gene expression; filamentous fungal cell; EST;
KW	expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW	Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW	culture condition; environmental stress; spore morphogenesis;
KW	metabolic pathway engineering; catabolic pathway engineering; ss.

OS	Aspergillus oryzae.	
XX		
XX	WO2000056762-A2.	
PN		
XX		
XX	28-SEP-2000.	
XX		
XX	22-MAR-2000; 2000WO-US07781.	
PF		
XX		
PR	22-MAR-1999; 99US-0273623.	
XX		
XX	(NOVO) NOVO NORDISK BIOTECH INC.	
PA	(NOVO) NOVO NORDISK AS.	
PA		
XX		
PI	Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;	
XX		
XX	WPI; 2000-594572/56.	
XX		
PT	Monitoring differential expression of genes in filamentous fungal cells	
PT	uses fluorescence-labeled nucleic acids isolated from the cells and a	
PT	substrate of expressed sequence tags -	
XX		
XX	Claim 88; Page 1893-1894; 3161pp; English.	
PS		

The present invention describes a method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be

monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene products to facilitate analysis of the results. AAF07478 to AAF11247 represents ESTs from *Fusarium venenatum*; AAF11248 to AAF1853 represents ESTs from *Aspergillus niger*; AAF1854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

SQ	Sequence	1358 BP;	258 A;	430 C;	348 G;	322 T;	0 other;
Query Match	Score	38.8;	DB 21;	Length	1358;		
Best Local Similarity	49.5%;	Pred.	No. 0.19;				
Matches	Conservative	0;	Mismatches	102;	Indels	0;	Gaps
QY	801	GATTATCGTTCATTATGCTCCCTTTCTTATTCGTATGGCTTGCCAGCGTGC CGGRACATAC	860				
Db	382	GATGACGGCAGCTAATGACCGCGTCTCGTCGCTGCGCATGGCACGAAGTGGCACTTAT	441				
QY	861	AGGACATATGATGCCGGGGAGGCGCCAGTAGTGGTGTGCTACGCAACGTTTTTAACCCGTGAAC	920				
Db	442	GACAAGCAGACTGCTACCGGAGGTAGCAAACGGTGCCACCATGAGATTCGCCCCCTGAGTCC	501				
QY	921	AGTGCGCGGATACAGTTATCTCGATATAAACCCCGTCGATGCTGTGTGCCAGCATCAAGAA	980				
Db	502	GACCACGGCGCCAACGCTGGTGTCTCAAGGCTCCCAGAGACTTCTCTGGAGCCTGTCAAGGCC	561				
QY	981	AAATACGGCTCCAGTATTTCT	1002				
Db	562	AAGTTCCTTGGATCACTACT	583				

RESULT 25

AAT89278	ID	AAT89278 standard; DNA; 4320 BP.
XX	AC	AAT89278;
XX	DT	19-FEB-1998 (first entry)
XX	DE	TRPP-F1 gene, including the promoter region.
XX	KW	TRPP-F1; tomato fruit expressed proline-rich protein gene; ovary; promoter; genetic parthenocarpy; auxin modulator; transgenic plant; ss.
XX	OS	Lycopersion esculentum.
XX	Key	Location/Qualifiers
FH	Promoter	1..2643 /*tag= a
FT	prim_transcript	2235..4320 /*tag= b
FT	CDS	2644..3684 /*tag= c
FT		
FN	WO9730165-A2.	
PN		
XX		
XX	21-AUG-1997.	
PD		
XX		
PF	13-FEB-1997; 97WO-IL00051.	
XX		
PR	14-FEB-1996; 96IL-0117139.	
XX		
PA	(ISRA) ISRAEL MIN AGRIC.	
XX		
PI	Barg R, Salts Y;	
XX		
DR	WPI; 1997-425040/39.	

	XX	Producing genetic parthenocarpy in plants - with sequence encoding auxin modulator controlled by promoter specific for the ovary
	PT	Claim 7; Fig 9; 41pp; English.
	XX	This sequence comprises a genomic clone of the tomato fruit expressed proline-rich protein gene PRP-F1, including the promoter region. A claimed method for the production of genetic parthenocarpy in plants comprises: providing a cassette including DNA coding for modulation of auxin effects in plants and a promoter specific for the ovary between anthesis and early fruit development; and introducing the cassette into a plant. Suitable promoters include the TPFP-F1 promoter and an alternative PRP-F1 promoter (see AR89279) that confers ovary and developing embryo specificity. The method allows parthenocarpy to be induced in especially tomato plants without the need to apply (synthetic) auxin or auxin inhibitor to each truss.
	CC	Sequence 4320 BP; 1450 A; 844 C; 585 G; 1441 T; 0 other;
	XX	Query Match 2.6%; Score 38.8; DB 18; Length 4320;
	SQ	Best Local Similarity 43.8%; Pred. No. 0.4;
		Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;
QY	424	CTTAGTATTTTTCGTGAGTATCATCCACGAATAGGTATATCCTGGTCATCAATA 483
DB	3628	CTTATTGATGTTGTGCCAGATATCACCCAAGACTTCAAGTGTCCTTCAACCTAATC 3687
QY	484	AAGTTGACTTTTGTATACAACATCGGAATTTCCCTTAATCCGGAGCTATTCGTATGATAA 543
DB	3688	AAGGTTTTCCACTTTTCTCACTTTCAAATTATTACTCACTCCTACTCAATTTATGTGTAC 3747
QY	544	AAAAAAGCTCTCCGTGTTCTGATTCCTTGCGCTATCGGGAGCTTTTCTACCGCTGTAG 603
DB	3748	AGTTGCATTTTCAAGTATTAGGCCCAAATTTCTTAGCTCGGAATTTTAAAATCTCTTT 3807
QY	604	CCGCTGTATAAAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTAACTCCTC 663
DB	3808	AAATATTTTGATTTTATCTACTTACTTATTACGTAGTTTTCATAGTATAAAATTTCAATTCATA 3867
QY	664	TGAGATTTACAGCCCTGAATCAAAATCCCCTGGGGGCGTGAATTCATGCCACCAGAT 723
DB	3868	TATGAATTCACGGTCAAAAATTTAAAAGTTTATTAAACCAATGCCACATAACTTAGAACAT 3927
QY	724	TTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTGCTGCAACACTTCCC 783
DB	3928	ACAAATATCTTTTGTATCAAGATTTGCCAATTCGTATACAATAAATCTTTAGCAAGTAATA 3987
QY	784	AGGATTTGTCGCCCTGCCGATATATGGT 809
DB	3988	TGTATACCAACATTTATGTAATATGAT 4013

RESULT 26	
ABL08713	Drosophila melanogaster expressed polynucleotide SEQ ID NO 20621.
XX	
XX	Drosophila; developmental biology; cell signalling; insecticide;
XX	pharmaceutical; gene; ss.
XX	
XX	Drosophila melanogaster.
XX	
XX	WO200171042-A2.
XX	
XX	27-SEP-2001.
XX	

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PPR 11-JUL-2000; 2000US-061415O.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX P-PSDB; ABB64610.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 20621; zlipp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 5591 BP; 1876 A; 1209 C; 1357 G; 1149 T; 0 other;
Query Match 2.5%; Score 37.8; DB 23; Length 5591;
Best Local Similarity 57.0%; Pred. No. 0.99;
Matches 69; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1329 ACATTGGTAAAGCACATGCTGCACCGCTCTCCTGAATAATGTATTCGCCGAGGCGCTCAT 1388
Db ||||| |||| |||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1358 ACATGTGGAGAAGCTGATGTTGGAGCGCGACCCTGGACCGGAGTGCCCGAACAACGAGC 1417
QY 1389 GTTGCACTCTGTGAGGAGGACGGACTGGGATGGAAAAATAAATGTGTACAGAAAACGGC 1448
Db ||||| |||| |||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1418 GCTGCAGCTGCAGAGAATCATCAACGAGCTAAAGCAGAAATCGTTGAATTGGAGTCGGC 1477
QY 1449 A 1449
Db I
 1478 A 1478

RESULT 27
ABL08655 standard; cdna; 6000 BP.
ID ABL08655 standard; cdna; 6000 BP.
XX
XX ABL08655;
XX
XX
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 20447.
XX DE
XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmacetical; gene; ss.
XX OS Drosophila melanogaster.
XX WO200171042-A2.
XX PN
XX PD
XX XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-061415O.
XX

PA (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB64552.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Claim 1; SEQ ID NO 20447; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 6000 BP; 2026 A; 1274 C; 1434 G; 1266 T; 0 other;
XX
XX Query Match 2.5%; Score 37.8; DB 23; Length 6000;
XX Best Local Similarity 57.0%; Pred. No. 1;
XX Matches 69; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
XX
XX QY 1329 ACATTGGTAAAGCACATGGTCAGCGCTCTCTCGAAATATGTTGGCGGAGGCGCTGAT 1388
XX DB 1767 ACATGTGGAGAGCTGATGGTGAGCGCGACCTGGACCGGAGGATGCCAGAACCGGC 1826
XX QY 1389 GGTGCACCTGTGGAGGAGGAGGAGCTGGGATGAAAAATATGTTGTACAGGAAACGGC 1448
XX DB 1827 GCTGCAGCTGCAGAAACATCAACGAGCTAAAGCAAGAAATCGTTGAATGGAGTCGGC 1886
XX QY 1449 A 1449
XX DB 1887 A 1887
XX
XX RESULT 28
XX AAX20256
XX ID AAX20256 standard; DNA; 14752 BP.
XX AC AAX20256;
XX AC AAX20256;
XX DT 04-MAY-1999 (first entry)
XX DE Borrelia burgdorferi polynucleotide sequence #9.
XX KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
XX KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
XX KW infection; diagnosis; characterisation; detection; ds.
XX OS Borrelia burgdorferi.
XX PN WO9858943-A1.
XX PD 30-DEC-1998.
XX PF 18-JUN-1998; 98WO-US12764.
XX PR 03-SEP-1997; 97US-0057483.
XX PR 20-JUN-1997; 97US-0050359.
XX PR 22-JUL-1997; 97US-0053344.
XX PR 22-JUL-1997; 97US-0053377.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.

XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
XX White OR;
XX WPI; 1999-081217/07.
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of infections, particularly Lyme disease
XX Claim 1; Page 893-901; 1128pp; English.
XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for
XX the detection, diagnosis, characterisation, prevention and therapy of
XX Bb infections, e.g. Lyme disease. They can also be used for the
XX production of biosynthetic products, e.g. enzymes. Borrelia belongs
XX to a family of motile, spiral-shaped bacteria called Spirochetes.
XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and
XX endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX Lyme disease.
XX Sequence 14752 BP; 5196 A; 1781 C; 1634 G; 6139 T; 2 other;
XX
XX Query Match 2.5%; Score 36.6; DB 20; Length 14752;
XX Best Local Similarity 57.4%; Pred. No. 4.5;
XX Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
XX
XX QY 426 TAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTCATCAATAA 485
XX DB 8104 TATTCCTTTTCTTTTATTAATGCTTCTGTAATAATAAAGCGGTGTTTCCAAAGA 8163
XX QY 486 GTTGACTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTCGTATGA 540
XX DB 8164 TTTTCTTTTGTATATAAATGCAAGCTTCGCTTATCTTTCGCTTAATGTTTTA 8218
XX
XX RESULT 29
XX AAF07750
XX ID AAF07750 standard; cDNA; 801 BP.
XX AC AAF07750;
XX DT 13-MAR-2001 (first entry)
XX DE Fusarium venenatum EST SEQ ID NO:273.
XX KW Multiple gene expression; filamentous fungal cell; EST;
XX KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
XX KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
XX KW culture condition; environmental stress; spore morphogenesis;
XX KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX OS Fusarium venenatum.
XX PN WO200056762-A2.
XX PD 28-SEP-2000.
XX PF 22-MAR-2000; 2000WO-US07781.
XX PR 22-MAR-1999; 99US-0273623.
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX (NOVO) NOVO NORDISK AS.
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX WPI; 2000-594572/56.
XX Monitoring differential expression of genes in filamentous fungal cells
XX uses fluorescence-labeled nucleic acids isolated from the cells and a
XX substrate of expressed sequence tags -

[illegible]

RESULT 35
ABL70580/c

QY 524 CGGACCTATTCGTATGATAAATAAACAATCTTCTGTTCTGATTCTTCTGGCGCTATCGGG 583
 DB 354 CTGCGCTGTTGCTATGATGCTCAATAATTCCTTGAGTCTGCGCCCTAAAGACTCTCGTGG 295
 QY 584 GAGCTTTTCTACCGCTAGCGCTGATGATAAATAAAGAGACT 623
 DB 294 GCCTGTTCCACCGAGACACAGACAGACAGCCGAGTGTCT 255

RESULT 39
 AAZ15175/C
 ID AAZ15175 standard; cDNA; 800 BP.
 AC AAZ15175;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 DE Human gene expression product cDNA sequence SEQ ID NO:2644.
 XX
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9938972-A2.
 XX
 PD 05-AUG-1999.
 XX
 PF 28-JAN-1999; 99WO-US01619.
 XX
 PR 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Gliese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX
 WPI; 1999-494092/41.
 XX
 DR Novel human genes and their expression products which are
 XX differentially expressed in different cell types
 PT
 PS Claim 1; Page 1286; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ1779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purposes, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for

CC peptide analogues and antagonists.
 XX
 SQ Sequence 800 BP; 255 A; 153 C; 158 G; 209 T; 25 other;
 Query Match 2.3%; Score 34.2; DB 20; Length 800;
 Best Local Similarity 53.4%; Pred. No. 4.4;
 Matches 63; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 611 TAAAAAGAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTCTCTGAGATT 670
 DB 793 TTAACCTAGCCAGGTTTCCCTGTCACCNAGACTGGAGTTAAGAACCTGANAAN 734
 QY 671 ACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATATGCCACAGATTCAA 728
 DB 733 AGCCCACTGTGCCTCAAAATTTCTGGGGCTCAAGTGTATCTCTGCTCAACNTCTAAA 676

RESULT 40
 AAS49611
 ID AAS49611 standard; DNA; 893 BP.
 AC AAS49611;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation inhibitory sequence #835.
 XX
 KW Antisense; ss; prokaryotic cellular proliferation;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 WPI; 2001-611495/70.

New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids -
 Claim 1; Seq ID No 2188; 511pp; English.
 The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence is an antisense oligonucleotide of the invention.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 893 BP; 273 A; 161 C; 119 G; 340 T; 0 other;

Query Match 2.3%; Score 34.2; DB 23; Length 893;
Best Local Similarity 49.7%; Pred. No. 4.7;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 483 AAGTTGACCTTTGTATACACATGCGAATTCCTTAATCCGGAGCTATTCGTATGATA 542
DB 30 AAGTTCAAAATGCTTCTGCATCTTGGAAATGCTTTTATCAGCACTCATCTAATA 89
QY 543 AAAAAACCTTCTCTGTCGATCTCTGCGCTATCCGGGAGCTTTTCTACCGTGTA 602
DB 90 ATAATTTTAAATTTATCTGCTTCTCAGTAGCACTTGTCCAGGATGAAAACTATACG 149
QY 603 GCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTAA 657
DB 150 TCGGCTCTAGTACAAATGATACCAAAATGATCATCGATAAGTCTTCGTTTAA 204

RESULT 41
AAS55352/C
ID AAS55352 standard; DNA; 927 BP.

AC AAS55352;
DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #1664.
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-PSDB; AAU37493.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 8989; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 927 BP; 347 A; 116 C; 169 G; 295 T; 0 other;
Query Match 2.3%; Score 34.2; DB 23; Length 927;
Best Local Similarity 49.7%; Pred. No. 4.8;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 483 AAGTTGACCTTTGTATACACATGCGAATTCCTTAATCCGGAGCTATTCGTATGATA 542
DB 715 AAGTTCAAAATGCTTCTGCATCTTGGAAATGCTTTTATCAGCACTCATCTAATA 656
QY 543 AAAAAACCTTCTCTGTCGATCTCTGCGCTATCCGGGAGCTTTTCTACCGTGTA 602
DB 655 ATAATTTTAAATTTATCTGCTTCTCAGTAGCACTTGTCCAGGATGAAAACTATACG 596
QY 603 GCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAACACTGGATTAA 657
DB 595 TCGGCTCTAGTACAAATGATACCAAAATGATCATCGATAAGTCTTCGTTTAA 541

RESULT 42
AAS52250/C
ID AAS52250 standard; DNA; 1164 BP.

AC AAS52250;

DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #667.

XX Antisense; ds; prokaryotic cellular proliferation gene;

KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-PSDB; AAU34391.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 4832; 51lpp; English.

XX PF 24-MAY-2001; 2001WO-US17495.
 XX PR 24-MAY-2000; 2000US-207138P.
 XX PA (UROC-) UROGENESYS INC.
 XX PI Challita-Eid PM, Hubert RS, Faris M, Afar DEH, Levin E;
 XX PI Mitchell SC, Jakobovits A;
 XX DR WPI; 2002-097642/13.
 XX PT New isolated 98p7C3-related homeodomain protein highly expressed in
 XX PT various cancers, useful in cancer vaccines and for generating immune
 XX PT response directed to 98p7C3 in mammal
 XX PS Example 1; Fig 1; 155pp; English.
 XX CC The invention relates to an isolated 98p7C3-related protein which is a
 XX CC homeodomain protein highly expressed in various cancers. Also include
 XX CC are polynucleotides encoding the protein or proteins 90% identical
 XX CC to 98p7C3, a pharmaceutical composition comprising the polynucleotides
 XX CC (including an expression vector comprising the 98p7C3 encoding
 XX CC polynucleotides) or a host cell transformed with the vector,
 XX CC an anti-98p7C3 antibody, a non-human transgenic animal expressing a
 XX CC 98p7C3 protein, methods of detecting the 98p7C3 protein or
 XX CC polynucleotides in a biological sample, monitoring the presence of
 XX CC cancer in an individual by detecting an elevated level of the 98p7C3
 XX CC protein or polynucleotides and a pharmaceutical composition comprising a
 XX CC modulator of 98p7C3 protein, or T cell/B cell epitopes derived
 XX CC from it, are useful in inducing an immune response (in mammal) to a
 XX CC 98p7C3 protein. Upon contact with a cytotoxic T cell (CTL) the immunogens
 XX CC induce the CTL (with its helper T cell) to kill an autologous cell
 XX CC expressing 98p7C3. The immunogen may be a nucleic acid encoding the
 XX CC protein or epitope. The antibody is useful for delivering a cytotoxic
 XX CC agent to a cell that expresses 98p7C3, by conjugating the cytotoxic
 XX CC to the antibody or its fragment that specifically binds to a 98p7C3
 XX CC epitope, and exposing the cell to the antibody-agent conjugate. The
 XX CC modulator is useful for treating a patient with a cancer that expresses
 XX CC 98p7C3 (e.g. prostate cancer, bladder cancer, kidney cancer, lung cancer,
 XX CC breast cancer, uterine cancer, cervical cancer, stomach cancer, rectal
 XX CC cancer and colon cancer), by administering to the patient a vector
 XX CC that comprises the modulator, such that the vector delivers a single
 XX CC chain monoclonal antibody coding sequence to the cancer cells and the
 XX CC encoded single chain antibody is expressed intracellularly in it.
 XX CC The gene for 98p7C3 is located on human chromosome 4q11-q12. The present
 XX CC sequence is oligonucleotide adapter or PCR primer used to isolate a cDNA
 XX CC hybridisation, SSH.
 XX SQ Sequence 188 BP; 55 A; 19 C; 15 G; 80 T; 19 other;
 Query Match 2.3%; Score 34; DB 24; Length 188;
 Best Local Similarity 57.8%; Pred. No. 2.1;
 Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 460 TAGGTATATCCCTGTCATCAATAAGTTGACCTTTGTATACAAATGCGAATTTCCCTT 519
 Db 113 TNAGTTTATTTTGTNTNAATAAANTGAACCTGCTTTGTCTCTAAATATNTTTTNAATN 54
 QY 520 ATCCGGAGCTATTCGTATGATATAAAAAA 549
 Db 53 AAACCTTAACCTTTTCTTGTGGAAAAA 24
 RESULT 45
 AAH94492
 ID AAH94492 standard; cDNA; 308 BP.
 XX AC AAH94492;
 XX DT 05-OCT-2001 (first entry)
 XX DE Human foetal cDNA, SEQ ID NO: 520.

DE XX Human foetal cDNA, SEQ ID NO: 1179.
 XX KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 XX KW nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antitense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation; expressed sequence tag; EST; ss.
 XX OS Homo sapiens.
 XX PN WO200155339-A2.
 XX PD 02-AUG-2001.
 XX PF 25-JAN-2001; 2001WO-US02723.
 XX PR 25-JAN-2000; 2000US-0491404.
 PR 15-SEP-2000; 2000US-0663870.
 PR 06-NOV-2000; 2000US-0707351.
 XX (HYSE-) HYSEQ INC.
 PA Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werhman T;
 XX WPI; 2001-465571/50.
 DR P-PSDB; AAM06817.
 XX Novel fetal proteins useful for the treatment and diagnosis of diseases
 XX associated with dysfunction of the protein e.g. cancers, immune
 XX disorders, growth disorders, thrombolytic disorders, nervous system
 XX disorders and inflammation -
 PS Example 3; Page 666; 715pp; English.
 CC The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are
 CC useful in the treatment and diagnosis of diseases such as cancers,
 CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
 CC disorders, nervous system disorders and inflammation. The present
 CC sequence is a full length cDNA which was assembled using expressed
 CC sequence tags (ESTs) found to be expressed in human foetal tissue
 CC cDNA libraries as seeds.
 XX SQ Sequence 308 BP; 59 A; 63 C; 46 G; 140 T; 0 other;
 Query Match 2.3%; Score 34; DB 22; Length 308;
 Best Local Similarity 53.8%; Pred. No. 2.8;
 Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 425 TTAGTATTTTCTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTGCAATATA 484
 Db 45 TAATTTCTTTCTTTCTAGGCTCTGTACTGATGAATATATATCTCTTTGGAGGTATCA 104
 QY 485 AGTTGACTTTTGTATACAAACATCGGAATTTCCCTTAATCCGGAGCTATTGTTATGATAA 544
 Db 105 TGTTCCTTCTTTTTCATGTTTCTTGTGCTTACATTAATATCTGCACATCTGATGTA 164
 QY 545 AAAAACTCTT 554-
 Db 165 AAGTCACCT 174
 RESULT 46
 AAH93991
 ID AAH93991 standard; cDNA; 316 BP.
 XX AC AAH93991;
 XX DT 05-OCT-2001 (first entry)
 XX DE Human foetal cDNA, SEQ ID NO: 520.

XX Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW neutropic; neuroprotective; thrombolytic; osteopathic; antinflammatory;
 KW gene therapy; antitense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation; expressed sequence tag; EST; ss.
 XX Homo sapiens.
 OS WO200155339-A2.
 PN 02-AUG-2001.
 PD 25-JAN-2001; 2001WO-US02723.
 PF 25-JAN-2000; 2000US-0491404.
 PR 15-SEP-2000; 2000US-0663870.
 PR 06-NOV-2000; 2000US-0707351.
 XX (HYSE-) HYSEQ INC.
 PA Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werhman T;
 XX WPI; 2001-465571/50.
 DR P-PSDB; AAM06316.
 XX Novel fetal proteins useful for the treatment and diagnosis of diseases
 PT associated with dysfunction of the protein e.g. cancers, immune
 PT disorders, growth disorders, thrombolytic disorders, nervous system
 PT disorders and inflammation -
 XX Claim 1; Page 367-368; 715pp; English.
 PS The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are
 CC useful in the treatment and diagnosis of diseases such as cancers,
 CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
 CC disorders, nervous system disorders and inflammation. The present
 CC sequence was assembled using an expressed sequence tag (EST) found
 CC to be expressed in human foetal tissue cDNA libraries as the seed.
 XX Sequence 316 BP; 61 A; 65 C; 50 G; 140 T; 0 other;
 SQ
 Query Match 2.3%; Score 34; DB 22; Length 316;
 Best Local Similarity 53.8%; Pred. No. 2.9;
 Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 425 TTAGTATTTTGTGCTGTAGTATCTATCCAGCAATAGTAGTATCCCTGTTGCATCAATAA 484
 Db 53 TAAATTTCTTTTCTTAGCGCTGTACTGATGATGATATATGTCCTTTGGAGGTATCA 112
 QY 485 AGTTGACATTTGTATACAAATGCGAATTTCCCTTAATCCGGAGCATTCGTATGATAAA 544
 Db 113 TGTTCCTTCTTTTTCATGTTCTTGTGTCCTTACATTAATATATGCAATCTGATGA 172
 QY 545 AAAAATCTTT 554
 Db 173 AAAGTCACTT 182
 RESULT 47
 AAS30450/c
 ID AAS30450 standard; DNA; 636 BP.
 XX AAS30450;
 AC
 XX 21-NOV-2001 (first entry)
 DT
 XX DNA encoding novel prostate gland antigen, Seq ID No 308.
 DE
 XX Human; neutropic; neuroprotective; cytostatic; antiparkinsonian;

KW antianemic; dermatological; immunosuppressive; antiinflammatory;
 KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;
 KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;
 KW prostatosis; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;
 KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;
 KW reproductive system disorder; autoimmune disorder; urinary system;
 KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;
 KW blood-related disorder; hyperproliferative disorder; respiratory;
 KW neurological disorder; endocrine disorder; inflammatory disorder;
 KW liver disorder; wound healing; food preservative; ds.
 XX Homo sapiens.
 OS WO200155447-A1.
 PN 02-AUG-2001.
 PD 17-JAN-2001; 2001WO-US01330.
 PF 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-MAR-2000; 2000US-0198123.
 PR 19-MAR-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-476223/51.
 XX
 XX Novel isolated prostate gland related polypeptide useful for diagnosis
 PT and treatment of disorders of prostate such as prostatodystonia,
 PT prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia
 PT
 XX
 PS Claim 1; SEQ ID No 308; 512pp; English.
 XX
 CC The invention relates to novel isolated prostate gland related nucleic
 CC acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,
 CC prognosis, prevention, and/or treatment of diseases and/or disorders of
 CC the prostate such as acute non-bacterial prostatitis, chronic non-
 CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,
 CC prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic
 CC hypertrophy or hyperplasia, and prostate neoplastic disorders, including
 CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and
 CC squamous cell carcinomas. (I), (II) and antibody to (II) are useful for
 CC diagnosing and treating reproductive system disorders (Paget's disease),
 CC autoimmune disorders (systemic lupus erythematosus, rheumatoid
 CC arthritis), blood-related disorders (sickle cell anaemia),
 CC hyperproliferative disorders, urinary system disorders
 CC (glomerulonephritis), cardiovascular disorders (arrhythmias),
 CC respiratory disorders, musculoskeletal system disorders, neural activity
 CC and neurological disorders (Alzheimer's disease and Parkinson's disease),
 CC endocrine disorders (Addison's disease), gastrointestinal disorders
 CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),
 CC pancreatic and gall bladder disorders, disorders of the large intestine,
 CC developmental and inherited disorders, diseases at the cellular level,
 CC and wound healing and epithelial cell proliferation. (I) or (II) is
 CC useful to prevent skin aging, for preventing hair loss, to maintain
 Query Match 2.3%; Score 34; DB 22; Length 636;
 Best Local Similarity 59.2%; Pred No. 4.4;
 Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 662 TCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTGATTCACACAG 721
 Db 573 TCTTGTGTTGCTCAGGCTGCTCAAACTCCTGCTAGCTGTTGATTTTGAGCCACTAA 514
 QY 722 ATTCAACAGCTGGATATGAGGCTCTGAAAAAGATA 759
 Db 513 GTAACAACAGCTTAACATGAAGTATATGAGTCAGCTA 476
 RESULT 48
 AAL06230/c
 ID AAL06230 standard; DNA; 636 BP.
 XX
 AC AAL06230;
 XX
 DT 21-NOV-2001 (first entry)
 XX Human reproductive system related antigen DNA SEQ ID NO: 8918.
 DE Human; reproductive system related antigen; reproductive system disorder;
 KW Human; reproductive system related antigen; gene therapy; ds.
 XX
 OS Homo sapiens.

XX WO200155320-A2.
PN 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01339.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX

PI	Rosen CA, Barash SC, Ruben SM;
XX	
DR	WPI; 2001-465570/50.
XX	
PT	Isolated nucleic acid molecule encoding a reproductive system antigen
PT	is used in preventing, treating or ameliorating a medical condition -
XX	
PS	Disclosure; SEQ ID NO 8918; 1297pp + Sequence Listing; English.
XX	
CC	The present invention provides the protein and coding sequences of a
CC	number of human reproductive system related antigens. These can be used
CC	in the prevention and treatment of reproductive system disorders,
CC	including cancer. The present sequence is a genomic sequence encoding a
CC	protein of the invention.
XX	
SQ	Sequence 636 BP; 153 A; 133 C; 129 G; 221 T; 0 other;
Query Match 2.3%; Score 34; DB 22; Length 636;	
Best Local Similarity 59.2%; Pred. No. 4.4;	
Matches	59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY	662 TC7GAGATTACAGCCCTGAATCAAACTCCTGGGGGTGATTTTGATTATGCCACCAG 721
Db	573 TC7TGTGTTCTCAGCGTGGTCTCAAACCTCTCTAGCCTGTTGTTTTGAGCCACTAA 514
QY	722 ATTTTCACAGCTGGATATGGAGGCTCTGAAAAAGATA 759
Db	513 GTAACACACCTTAACATGAAGTATATGAAGTCAGCTA 476

RESULT 49	
AAQ12000/c	
ID	AAQ12000 standard; cDNA; 1844 BP.
XX	
XX	AAQ12000;
XX	
XX	
DT	25-SEP-1991 (first entry)
XX	
DE	Maize nitrite reductase.
XX	
XX	
KW	Nitrite reductase; nitrate; maize; spinach; probe; NR; ss.
XX	
OS	Zea mays L.
XX	
Key	Location/Qualifiers
CDS	9..1709
FT	/*tag= a
FT	/product= NR
XX	
XX	
PN	US4997930-A.
XX	
XX	
PD	05-MAR-1991.
XX	
XX	
PF	16-MAR-1989; 89US-0324154.
XX	
XX	
PR	16-MAR-1989; 89US-0324154.
XX	
XX	
PA	(CIBA) CIBA-GEIGY CORP.
XX	
PI	Lahnern KN, Rothstein SJ;
XX	
XX	
DR	WPI; 1991-206821/28.
DR	P-PSDB; AAR12517.
XX	
XX	
PT	Cloned DNA sequence encoding maize nitrite reductase - useful for
PT	determin. of NR number and study of NR mRNA regulation and nitrate
PT	assimilation.
XX	
XX	Claim 2; Page 9-10; 10pp; English.
XX	
XX	
CC	The sequence may be used to determine the number of NR genes per
CC	maize genome and to study the effect of various environmental
CC	conditions, incl. the presence of nitrate on the level of NR mRNA of

CC	maize plants grown under selected conditions. The cDNA may also be used to study NR mRNA regulation in maize. The cDNA is obtd. using lambda gtl1 as the cloning vector and a spinach NR cDNA as heterologous probe. The cDNA clone (pCIB808) is 66% homologous at the nucleotide level with the spinach NR gene, however the maize DNA has a high G/C content (69.5%) as compared to that of the spinach gene (46%) and is 75% homologous at the amino acid level.
XX	
SQ	Sequence 1844 BP; 338 A; 582 C; 655 G; 269 T; 0 other;
	Query Match 2.3%; Score 34; DB 12; Length 1844;
	Best Local Similarity 56.1%; Pred. No. 8.6;
	Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY	169 GAAGCATCAGAAGGCGGGGAACCTCCGTCCGCCAGTGAAACCGTGCACACTCCGGCA 228
Db	584 GACCGCGGCGAGCGGGTTGCCGACGGGTTGCCATGCCGCTTCGAGGCT 525
QY	229 GTACATCGCGCGCGGTGTATACC GGCGAAGATGGTCGAACCTCCCCTCCGT 282
Db	524 GGTGAGCGCACGCGCGGAGGCGGTCCAGGATGCCGGGACGTC CGGGAGCGT 471
RESULT 50	
AAK65489	
ID	AAK65489 standard; DNA; 16605 BP.
XX	
AC	AAK65489;
XX	
DT	06-NOV-2001 (first entry)
DE	
XX	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20301.
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW	cytostatic; gene therapy; vaccine; metastasis; ds.
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.
PF	
XX	17-JAN-2001; 2001WO-US01354.
PR	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225267.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.

Db 14753 GATATAAATACATAAAATTCACCTCTTCTATATACCATTTCCATGCTTTATGACAAGTG 14812
 Qy 588 TTTTCTACCGCTGTAGCGCTGATAAATAAGAGAGACTCAAAATTTCTACTATCCAGAAACA 647
 Db 14813 TGTCTAAGTTCTATATACTACTACCACAGTTTGAGACTTAAATTTCTACTATCTCAAAAAG 14872
 Qy 648 CTGGATTTAACTCCTCTG 665
 Db 14873 TTTCTTAGCCACTTCAG 14890

Search completed: June 18, 2003, 02:09:14
 Job time : 270.397 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 13:22:38 ; Search time 198.603 Seconds
(without alignments)
13391.616 Million cell updates/sec

Title: US-09-674-277-2

Perfect score: 1181
Sequence: 1 ctgcaggagatggaataaa.....ttttactttttctctgcag 1181

Scoring table: IDENTIFY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

N_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1181	100.0	1181	21	AAZ36102
2	159.4	13.5	1642	15	AAQ55982
3	64.2	5.4	10801	22	AAQ52323
4	57.2	4.8	4403765	22	AAI99683
5	57.2	4.8	4411529	22	AAI99682
6	49	4.1	1135	24	ABK74404
7	48.6	4.1	783	24	ABK78827
8	46.2	3.9	783	24	ABK78820
9	41.6	3.5	1227	24	ABK74480

C 10	40.8	3.5	1098	24	ABQ90082	M. capsulatus gene
C 11	40.6	3.4	14245	21	AAZ45259	DNA encoding enzym
C 12	40.6	3.4	14245	21	AAZ30356	Nucleotide sequenc
C 13	40.6	3.4	130480	22	AAF25833	R. marinus bacteri
C 14	39.8	3.4	4403765	22	AAI99683	Mycobacterium tube
C 15	39.8	3.4	4411529	22	AAI99682	Mycobacterium tube
C 16	39.6	3.4	265118	22	AAH41227	Mycobacterium tube
C 17	39.2	3.3	5059	20	AAH84332	Pyrococcus abyssi
C 18	39.2	3.3	349980	22	AAH41226	Stealth virus nucl
C 19	38.4	3.3	1110	24	ABQ90189	Pyrococcus abyssi
C 20	37	3.1	1108	24	ABK74439	M. capsulatus gene
C 21	36.6	3.1	386	23	ABV47264	Bacillus lichenifo
C 22	36.6	3.1	654	24	AAZ22921	Human prostate exp
C 23	36.6	3.1	1614	21	AAZ94534	Human soluble zaip
C 24	36.2	3.1	598	22	ABA60358	Human cytochrome rec
C 25	36.2	3.1	598	22	AAK34523	Human bone marrow
C 26	36.2	3.1	598	22	AAI40240	Human foetal liver
C 27	36.2	3.1	741	24	AAZ22970	Probe #8926 used t
C 28	36.2	3.1	1029	24	AAQ90206	Human soluble zaip
C 29	36.2	3.1	1704	22	AAQ06409	M. capsulatus gene
C 30	36.2	3.1	10732	21	AAI0594	Human DNAX cytokin
C 31	36	3.0	650	23	ABV03347	Gene encoding a su
C 32	36	3.0	1427	24	ABK73466	Human prostate exp
C 33	36	3.0	1859	22	AAH29949	Bacillus lichenifo
C 34	35.8	3.0	426	23	AA565293	S cerevisiae apopt
C 35	35.8	3.0	488	23	ABV48872	DNA encoding novel
C 36	35.8	3.0	5059	20	AAH84332	Human prostate exp
C 37	35.8	3.0	5666	22	ABA19027	Stealth virus nucl
C 38	35.8	3.0	5666	22	AAK75576	Human nervous syst
C 39	35.6	3.0	1614	22	AAQ06413	Human immune/haema
C 40	35.6	3.0	2163	21	AAZ61263	Human DNAX cytokin
C 41	35.4	3.0	61313	23	AA559545	DNA encoding a suc
C 42	35	3.0	1135	24	ABK74411	Propionibacterium
C 43	34.6	2.9	355	21	AAZ1988	Bacillus lichenifo
C 44	34.2	2.9	21410	22	AAK1290	Human secreted pro
C 45	34.2	2.9	74962	22	AAK15256	Human immune/haema
C 46	34	2.9	645	23	AA570590	Human phosphatase
C 47	34	2.9	1586	21	AAK43719	DNA encoding novel
C 48	33.8	2.9	333	22	AAK60217	zea mays DNA fragm
C 49	33.8	2.9	474	23	ABV48867	Human immune/haema
C 50	33.8	2.9	1647	13	AAQ21621	Human prostate exp
C 51	33.8	2.9	1712	24	ABN96850	Human Tryptophan-2
C 52	33.8	2.9	3442	13	AAQ21623	Gene #3348 used to
C 53	33.6	2.8	471	23	ABV48730	Human Tryptophan O
C 54	33.6	2.8	15463	22	AAK86350	Human prostate exp
C 55	33.4	2.8	484	23	ABV46504	Human immune/haema
C 56	33.4	2.8	494	23	ABV46580	Human prostate exp
C 57	33.4	2.8	2904	23	AA580187	Human prostate exp
C 58	33.2	2.8	516	23	ABV48886	DNA encoding novel
C 59	33.2	2.8	587	23	ABV48840	Human prostate exp
C 60	33.2	2.8	4590	22	AAH24065	Human prostate exp
C 61	33.2	2.8	5574	24	AA563339	Yeast AOD9604-asso
C 62	33	2.8	373	23	ABV47287	Chemically pretrea
C 63	33	2.8	492	23	ABV47754	Human prostate exp
C 64	33	2.8	3147	20	AAV81786	Human prostate exp
C 65	33	2.8	5570	20	AAV81789	Granulocytic Ehrli
C 66	32.8	2.8	366	20	AAZ21095	Polynucleotide seq
C 67	32.8	2.8	501	23	ABV48945	Human prostate exp
C 68	32.6	2.8	282	23	ABV47775	Human prostate exp
C 69	32.6	2.8	460	23	ABV46381	Human prostate exp
C 70	32.6	2.8	460	23	ABV47995	Human prostate exp
C 71	32.6	2.8	467	23	ABV48839	Human prostate exp
C 72	32.6	2.8	491	23	ABV47064	Human prostate exp
C 73	32.6	2.8	497	23	ABV46335	Human prostate exp
C 74	32.6	2.8	516	23	ABV48938	Human prostate exp
C 75	32.6	2.8	587	23	ABV46370	Human prostate exp
C 76	32.6	2.8	2426	20	AAZ10410	DNA encoding human
C 77	32.4	2.7	427	23	ABV32233	Human prostate exp
C 78	32.4	2.7	427	23	ABV41164	Human prostate exp
C 79	32.4	2.7	427	23	ABV41178	Human prostate exp
C 80	32.4	2.7	599	23	ABV48955	Human prostate exp
C 81	32.4	2.7	1122	20	AAH84119	Human prostate exp
C 82	32.4	2.7	11649	22	AAI03359	Pseudomonas aerugi
						Human reproductive

83	32.4	2.7	11649	22	AAS28910	Human immunoglobul
c 84	32.4	2.7	11649	22	AAK89078	Human digestive sy
c 85	32.4	2.7	11650	22	AAK89077	Human digestive sy
c 86	32.4	2.7	12441	20	AAK06750	Salmonella enteric
c 87	32.2	2.7	561	23	ABV47736	Human prostate exp
c 88	32.2	2.7	564	23	ABV48561	Human prostate exp
c 89	32.2	2.7	1254	22	AAH65416	C glutamicum codin
c 90	32.2	2.7	1272	24	ABQ90218	M. capsulatus gene
c 91	32.2	2.7	1689	24	ABK63827	Rat sequence diffe
c 92	32.2	2.7	3519	16	AAO94352	BHVI gi glycoprote
c 93	32.2	2.7	4383	23	ABL07815	Drosophila melanog
c 94	32.2	2.7	7178	24	ABK84453	Human cdna differe
c 95	32.2	2.7	7178	24	ABL61941	Colon adenocarcino
c 96	32.2	2.7	7367	24	ABK84452	Human cdna differe
c 97	32.2	2.7	9721	23	ABL07586	Drosophila melanog
c 98	32.2	2.7	349980	22	AAH68525	C glutamicum codin
c 99	32	2.7	435	23	ABV49147	Human prostate exp
100	32	2.7	487	23	ABV32296	Human prostate exp
101	32	2.7	487	23	ABV41127	Human prostate exp
102	32	2.7	487	23	ABV44673	Human prostate exp
103	32	2.7	493	23	ABV48448	Human prostate exp
104	32	2.7	601	23	ABV48230	Human prostate exp
c 105	32	2.7	1311	23	AA552027	Staphylococcus aur
c 106	32	2.7	1563	23	AA554470	Staphylococcus aur
c 107	32	2.7	4590	22	AAH24065	Yeast AOP3604-asso
108	32	2.7	50000	21	AAK96367	Polymorphic repeat
109	31.8	2.7	400	23	ABV16382	Human prostate exp
110	31.8	2.7	456	23	ABV47610	Human prostate exp
111	31.8	2.7	479	23	ABV49156	Human prostate exp
112	31.8	2.7	485	23	ABV48591	Human prostate exp
113	31.8	2.7	487	23	ABV48156	Human prostate exp
114	31.8	2.7	502	23	ABV48780	Human prostate exp
115	31.8	2.7	505	23	ABV48972	Human prostate exp
116	31.8	2.7	509	23	ABV49208	Human prostate exp
117	31.8	2.7	509	23	ABV49696	Human prostate exp
118	31.8	2.7	575	23	ABV48521	Human prostate exp
119	31.8	2.7	597	23	ABV46464	Human prostate exp
120	31.8	2.7	619	23	ABV46740	Human prostate exp
c 121	31.8	2.7	1146	24	AAJ28444	Bacillus subtilis
c 122	31.8	2.7	3153	22	AAE79464	Murine brevidin cd
123	31.8	2.7	7555	23	ABL10108	Drosophila melanog
124	31.6	2.7	276	23	ABV49839	Human prostate exp
125	31.6	2.7	362	23	ABV49714	Human prostate exp
126	31.6	2.7	410	23	ABV44750	Human prostate exp
127	31.6	2.7	411	23	ABV49771	Human prostate exp
128	31.6	2.7	437	23	ABV48648	Human prostate exp
129	31.6	2.7	449	23	ABV47740	Human prostate exp
130	31.6	2.7	458	23	ABV49161	Human prostate exp
131	31.6	2.7	460	23	ABV46085	Human prostate exp
132	31.6	2.7	460	23	ABV48897	Human prostate exp
133	31.6	2.7	462	23	ABV46017	Human prostate exp
134	31.6	2.7	462	23	ABV49236	Human prostate exp
135	31.6	2.7	467	23	ABV48423	Human prostate exp
136	31.6	2.7	469	23	ABV49626	Human prostate exp
137	31.6	2.7	471	23	ABV47053	Human prostate exp
138	31.6	2.7	474	23	ABV47863	Human prostate exp
139	31.6	2.7	474	23	ABV49355	Human prostate exp
140	31.6	2.7	475	23	ABV48575	Human prostate exp
141	31.6	2.7	476	23	ABV48002	Human prostate exp
142	31.6	2.7	479	23	ABV48777	Human prostate exp
143	31.6	2.7	479	23	ABV49527	Human prostate exp
144	31.6	2.7	481	23	ABV47790	Human prostate exp
145	31.6	2.7	481	23	ABV48948	Human prostate exp
146	31.6	2.7	485	23	ABV46908	Human prostate exp
147	31.6	2.7	490	23	ABV47641	Human prostate exp
148	31.6	2.7	490	23	ABV47702	Human prostate exp
149	31.6	2.7	491	23	ABV49046	Human prostate exp
150	31.6	2.7	492	23	ABV49119	Human prostate exp

ALIGNMENTS

RESULT 1

AAZ36102

AAZ36102 standard; DNA; 1181 BP.

XX

AC

AAZ36102;

XX

DT

11-FEB-2000 (first entry)

XX

DE

Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX

KW

Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

XX

KW

enterohemolysine; ehly; intimin; eae; virk gene; E. coli O157:H7; ds.

XX

OS

Escherichia coli.

XX

PN

W09955908-A2.

XX

PD

04-NOV-1999.

XX

PF

27-APR-1999; 99WO-FR01000.

XX

PR

28-APR-1998; 98PR-0005329.

XX

PA

(SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX

PI

Frechon DTM, Laure FC, Thierry D;

XX

DR

WPI; 2000-013443/01.

XX

PT

New nucleic acid containing sequences specific to enterohemorrhagic

XX

PT

Escherichia coli, particularly serotype O157:H7, used for detecting

XX

PS

these bacteria in food

XX

XX

Claim 1; Fig 2; 48pp; French.

XX

CC

The present sequence is specific to enterohemorrhagic Escherichia coli

CC

(EHEC). The sequence associated with the presence of virulence

CC

factors enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570

CC

also have 88% homology with the virk gene which codes for virulence

CC

proteins of Shigella flexneri. The present sequence is of plasmid origin.

CC

Fragments of the present sequence are used, as probes and primers, for

CC

detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),

CC

in human or animal samples, foods or the environment. The fragments are

CC

also useful for epidemiological studies.

XX

SQ

Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 other;

Query Match 100.0%; Score 1181; DB 21; Length 1181;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1

CTGCAGGAGATGGAAGAAAGCCAAATATAAATTTGCCATCCAGCGCGCTCCAGCTG 60

Db

1

CTGCAGGAGATGGAAGAAAGCCAAATATAAATTTGCCATCCAGCGCGCTCCAGCTG 60

QY

61

AAAGTAGGCTGTCTGTCGGGTATTTAATGCATGACCGTCCCGGTATTTAAACAATG 120

Db

61

AAAGTAGGCTGTCTGTCGGGTATTTAATGCATGACCGTCCCGGTATTTAAACAATG 120

QY

121

TGATAAATTACTCCGTACCGGAAACCCGCTGAAACAAATTCGGGTGAAAGAGATCC 180

Db

121

TGATAAATTACTCCGTACCGGAAACCCGCTGAAACAAATTCGGGTGAAAGAGATCC 180

QY

181

GGCGTTATCTGTGTCATTTCCCTTAGCTGACTGACCCAGACACAATGATCTGTGCGG 240

Db

181

GGCGTTATCTGTGTCATTTCCCTTAGCTGACTGACCCAGACACAATGATCTGTGCGG 240

QY

241

TTCTGTATATCAAAACCGGTACTCAATATCTTCTGGCGTGGCTGCATCCGGA 300

Db

241

TTCTGTATATCAAAACCGGTACTCAATATCTTCTGGCGTGGCTGCATCCGGA 300

QY

301

AGCGTTCCGCTCGGATATAAAATTCGCGAGTCCGCGGTCCATGCAGACATCCCCCA 360

XX DE Plasmid pAN240 carrying the B. subtilis PanBCD operon.
 XX DE PanBCD operon: pantothenate biosynthesis; cyclic; circular; pAN240;
 KW vitamin B5; nutritional supplement; panto-compound; pantoate; ds.
 XX OS Chimeric - Bacteriophage SP01.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Bacillus subtilis.
 XX PN WO200121772-A2.
 XX PD 29-MAR-2001.
 XX PF 21-SEP-2000; 2000WO-US25993.
 XX PR 21-SEP-1999; 99US-0400494.
 PR 07-JUN-2000; 2000US-0210072.
 PR 28-JUL-2000; 2000US-0221836.
 PR 24-AUG-2000; 2000US-0227860.
 XX PA (OMNI-) OMNIGENE BIOPRODUCTS.
 XX PI Yocum RR, Patterson TA, Hermann T, Pero JG;
 XX WPI; 2001-218644/22.
 XX PT New recombinant microorganism which overexpress a Bacillus subtilis
 PT pantothenate biosynthetic enzyme, useful for the high yield production
 PT of panto-compounds such as pantothenate and pantoate -
 XX Example 1; Page 232-238; 292pp; English.
 XX The sequence, pAN240, is a plasmid carrying the B. subtilis PanBCD operon
 CC which encodes three enzymes of the pantothenate biosynthetic pathway
 CC (PanB, C and D). Pantothenate, also known as vitamin B5, is used as a
 CC nutritional supplement in mammals and humans. The invention concerns
 CC methods of producing recombinant microorganisms overexpressing at least
 CC one B. subtilis pantothenate biosynthetic enzyme. The microorganisms and
 CC methods of producing them are useful for producing a panto-compound such
 CC as pantothenate or pantoate, which is a nutritional requirement for
 CC livestock and humans. The methods are also useful for the identification
 CC of pantothenate kinase modulators. Panto-compounds are produced at a
 CC significantly higher yield than prior art methods and can be produced
 CC independent of the need to feed precursors which decreases expense.
 XX Sequence 10801 BP; 2797 A; 2486 C; 2151 G; 3367 T; 0 other;
 SQ
 Query Match 5.4%; Score 64.2; DB 22; Length 10801;
 Best Local Similarity 54.4%; Pred. No. 1.2e-09;
 Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 335 GCCGTCCATGCAGACATCCCCCAGGGTAACAGCGTCCCTGTACATCTTCTGAAT 394
 DB 6490 GCCTGTCGCGCGTAACATACCGAGATCCACCACCAAAATCCGCTCACATGTTTTAT 6549
 QY 395 GACATCAGGATACCGCGCTTCCTACCTGGGATTAACGGGACGCGGAGACTGACGGTTC 454
 DB 6550 AACCTCAGGATACCGCAATGTTTGTTCATTAACAGGCACTCCGCAAGCATCGCTTC 6609
 QY 455 ACCGATACCATACCAACGCTTCATTTTCGAGGATGACACCACTGCGCAATCCG 514
 DB 6610 AAGCAGGACAGGCGCAAGCTTTCTTTTCAGATGACGAGCTTCAATCGCTTAATAGA 6669
 QY 515 GTAGACCGGTACAGCTGGGAAAGGCGGACCTGCCATTAAACATCTCCGTCATTC 571
 DB 6670 ATAAAGATCTTCACACGCTTCTTATTTCCAGCATTAAGACTTGGTCTTCCAGGCC 6726
 RESULT 4
 AAI99683
 ID AAI99683 standard; DNA; 4403765 BP.
 XX

AC AAI99683;
 XX 15-JAN-2002 (first entry)
 DT Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
 DE Mycobacterium tuberculosis
 XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
 KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
 XX Mycobacterium tuberculosis.
 OS US6294328-B1.
 XX PN 25-SEP-2001.
 XX PD 24-JUN-1998; 98US-0103840.
 XX PF 24-JUN-1998; 98US-0103840.
 PR (GENO-) INST GENOMIC RES.
 XX PA Fleischmann RD, White OR, Fraser CM, Venter JC;
 XX WPI; 2001-647261/74.
 XX DR Evaluating strain variation of Mycobacterium tuberculosis, comprises
 XX determining the nucleotide sequence of the strain at positions in the
 CC genome corresponding to positions where M. tuberculosis strains CDC
 CC 1551 and H37Rv differ -
 CC Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
 CC The invention relates to evaluating strain variation within and between
 CC different populations of the tuberculosis bacterial pathogen,
 CC Mycobacterium tuberculosis or related Mycobacterium by determining the
 CC nucleotide sequence of the first strain at positions in the complete
 CC sequence of the genome that correspond to positions that differ in the
 CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
 CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
 CC M. tuberculosis and has valuable application in the fields of
 CC tuberculosis genetics, epidemiology, patient treatment and epidemic
 CC monitoring.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
 XX SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
 Query Match 4.8%; Score 57.2; DB 22; Length 4403765;
 Best Local Similarity 54.2%; Pred. No. 8.3e-06;
 Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 272 TTCTCTGGCGCTGGCTGCCATCATCCGGAAGGTTCCGGTCGGGATATAAAATCGCGAG 331
 DB 2447373 TTCTCGCGCGGGCGGCCCATCCGACGCGCCGATCCGATCAGCACTCGGCCAC 2447432
 QY 332 TGGCGCGGTCATGACAGACATCCCGACAGGGTAACAGCGTCCCTGTCTCATTTCTTG 391
 DB 2447433 GCGCTCGGCCACCGCGTCCACCGACCTACCGTTCGACACTAGCCCACTTTGTGCTG 2447492
 QY 392 AATGACATCAGGATCCCGCCCTCTCACTGGCGATAACGGCACCGGAGACTGACGC 451
 DB 2447493 CACCGTTCCGGCGCTCCCGCAGATTCGGCGGATTAACGGCACCGCGCGGAGGC 2447552
 QY 452 TTCAGCAGTACCATACCAACAGCTTCAATTTTC 485
 DB 2447553 TTCAGGAAACCATGATGCCAAGCCCTCGACGTCC 2447586
 RESULT 5
 AAI99682
 ID AAI99682 standard; DNA; 4411529 BP.
 XX


```

AC  AAI99682;
XX
DT  15-JAN-2002 (first entry)
XX
DE  Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
DE  Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX  variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS  Mycobacterium tuberculosis.
XX
PN  US6294328-B1.
XX
PD  25-SEP-2001.
XX
PF  24-JUN-1998; 98US-0103840.
XX
PR  24-JUN-1998; 98US-0103840.
XX
PA  (GENO-) INST GENOMIC RES.
XX
PI  Fleischmann RD, White OR, Fraser CM, Venter JC;
XX  WPI; 2001-647261/74.
XX
DR  Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX  determining the nucleotide sequence of the strain at positions in the
XX  genome corresponding to positions where M. tuberculosis strains CDC
XX  1551 and H37Rv differ.
XX
PS  Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC  The invention relates to evaluating strain variation within and between
CC  different populations of the tuberculosis bacterial pathogen,
CC  Mycobacterium tuberculosis or related Mycobacterium by determining the
CC  nucleotide sequence of the first strain at positions in the complete
CC  sequence of the genome that correspond to positions that differ in the
CC  nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC  H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC  M. tuberculosis and has valuable application in the fields of
CC  tuberculosis genetics, epidemiology, patient treatment and epidemic
CC  monitoring.
CC
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from USPTO
CC  at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ  Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Query Match      4.8%; Score 57.2; DB 22; Length 4411529;
Best Local Similarity 54.2%; Pred. No. 8.3e-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY  272 TTCTCTGGCGTGGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGGCGAG 331
DB  2450072 TTCTCGCGCGGGCCCGCCATCGGAGCGCCGATCCCGATCGATCAGCACTCGGCCAC 2450131
QY  332 TGGCGCGGTCCATGACAGACATATCCCGCCACGGGTAAACGGTCCCTGTGTACATTTCTGTG 391
DB  2450132 GCGCTCGGCCACCGCGTCCACCGACCTACCTCGGACCACTAGCCGCTTGTGTGTG 2450191
QY  392 AATGACATCAGGATCCCGCGCTCTCACTGGGATAACGGCACCGGAGACTGACGC 451
DB  2450192 CACCGTTTCGGCGTCCCGCGGAGATTCGCGGAGTATCCGGCACCGCGCGGAGGC 2450251
QY  452 TTCAGCCAGTACCATCAACAAAGCTTCATTTTCC 485
DB  2450252 TTCGAGGACAGATGCCCAAGCCCTCGAGCTCC 2450285

RESULT 6
ID  ABR74404/C
ID  ABR74404 standard; DNA; 1135 BP.
XX

```

```

AC  ABK74404;
XX
DT  13-AUG-2002 (first entry)
XX
DE  Bacillus licheniformis genomic sequence tag (GST) #1695.
XX
DE  Differential gene expression; genomic sequenced tag; GST;
XX  altered culture condition; environmental stress;
XX  physiological provocation; ds.
XX
OS  Bacillus licheniformis.
XX
PN  WO200229113-A2.
XX
PD  11-APR-2002.
XX
PF  05-OCT-2001; 2001WO-US31437.
XX
PR  06-OCT-2000; 2000US-0680598.
XX  27-MAR-2001; 2001US-279526P.
XX
PA  (NOVO ) NOVOZYMES BIOTECH INC.
XX  (NOVO ) NOVOZYMES AS.
XX
PI  Berka R, Clausen IG;
XX
PI  WPI; 2002-416684/44.
XX
DR  Monitoring differential expression of several genes in first Bacillus
XX  cell relative to expression of same genes in one or more second
XX  Bacillus cells, by using substrate containing Bacillus genomic
XX  sequenced tag array.
XX
PS  Claim 4; SEQ ID NO 1695; 200pp; English.
XX
CC  The invention describes a method of monitoring differential expression of
CC  genes in a first Bacillus cell relative to expression of the genes in
CC  other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC  isolated from Bacillus cells to a substrate containing array of Bacillus
CC  genomic sequenced tags (GST), examining the array, and determining
CC  relative gene expression by an observed hybridisation reporter signal of
CC  a spot in the array. The method is useful for measuring the expression of
CC  genes in a first Bacillus cell relative to expression of the same genes
CC  in one or more second Bacillus cells. The method is useful for monitoring
CC  global expression of several genes from a Bacillus cell, discovering new
CC  genes, identifying possible functions of unknown open reading frames and
CC  monitoring gene copy number variation and stability. Monitoring changes
CC  in expression of genes may be used to provide a representation of the way
CC  in which Bacillus cells adapt to changes in culture conditions,
CC  environmental stress or other physiological provocation. Extensive
CC  follow-up characterisation is unnecessary, when one spot on an array
CC  equals one gene or one open reading frame, since sequence information is
CC  available. This sequence represents a genomic sequence tag (GST) used in
CC  the method of the invention.
CC
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at
CC  ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 1135 BP; 327 A; 251 C; 278 G; 279 T; 0 other;

Query Match      4.1%; Score 49; DB 24; Length 1135;
Best Local Similarity 50.2%; Pred. No. 3.5e-05;
Matches 121; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY  351 ACATCCCCCAGGGTAACAGCGTCCCTGTGCACATTTCTTGTGATGATCAGATCAGGATCCCG 410
DB  977 ATATCTCCGACATCGACCAAGTAACCGGATATCCGTTTTTGTGATCAGCTCGGGAATCCCC 918
QY  411 CCGGTCTACTGGCGATAACGGGACGCCGAGACTGACCTTTCAGCTCAGCTACCATACCA 470
DB  917 CCGATATTGTGCGGATGACGAGCGCCGAGCCATCGCTTCAAGCAGCAGCTAAACCG 858

```


CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 783 BP; 224 A; 150 C; 204 G; 203 T; 2 other;

XX Query Match 3.9%; Score 46.2; DB 24; Length 783;

XX Best Local Similarity 59.5%; Pred. No. 0.00024;

XX Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 354 TCCCCCAGGGGTAAACGGTCCCTGTGACATTTCTGTAATGACATCAGGATCCGCC 413

DB 633 TTCCCAAGGGGGCATAAATAAACAGTTACTTCTTCAACAACTTCCGGGATGCCACCA 574

QY 414 GTCTCACTGGCGATAACGGCGACCGGAGACTGACGCTTCACCGAGTACCAACCAAC 473

DB 573 ATTTTGTGCGGATCAGACAGTACCCCGCATGCCATTCACAGCGCCACTAAACCAAG 514

QY 474 GCTTCATTTC 484

DB 513 CTTCCTTTTC 503

RESULT 9

ABK74480/c

ID ABK74480 standard; DNA; 1227 BP.

XX ABK74480;

XX 13-AUG-2002 (first entry)

XX Bacillus licheniformis genomic sequence tag (GST) #1771.

XX Differential gene expression; genomic sequenced tag; GST;

XX altered culture condition; environmental stress;

XX physiological provocation; ds.

XX Bacillus licheniformis.

XX WO200229113-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US31437.

XX 06-OCT-2000; 2000US-0680598.

XX 27-MAR-2001; 2001US-279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX (NOVO) NOVOZYMES AS.

XX Berka R, Clausen IG;

XX WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus

XX cell relative to expression of same genes in one or more second

PT

PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array

XX Claim 4; SEQ ID NO 1771; 200pp; English.

XX The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1227 BP; 368 A; 236 C; 300 G; 323 T; 0 other;

XX Query Match 3.5%; Score 41.6; DB 24; Length 1227;

XX Best Local Similarity 62.5%; Pred. No. 0.01;

XX Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 352 CATCCCCCAGGGTAAACGGTCCCTGTGACATTTCTGTAATGACATCAGGATCCGC 411

DB 1072 CATCTCTTTAGCGGTACCATACCGGTCAATATGTTCAACGATTCGGGAAGTCCCC 1013

QY 412 CGCTCTCACTGGCGATAACGGCGACCGGAGACTGACGCTTCA 455

DB 1012 CTGCATCGCTGACGATAAATAGGAATCCGGCTAATTCGCGTTCA 969.

RESULT 10

ABQ90082/c

ID ABQ90082 standard; DNA; 1098 BP.

XX ABQ90082;

XX 01-OCT-2002 (first entry)

XX M. capsulatus gene #67 for DNA array.

XX Micro array; gene; ds; differential expression; gene expression.

XX Methylococcus capsulatus.

XX WO200255655-A2.

XX 18-JUL-2002.

XX 14-JAN-2002; 2002WO-NO00019.

XX 12-JAN-2001; 2001NO-0000235.

XX 12-JAN-2001; 2001NO-0000239.

XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.

XX (TIGR-) TIGR.

XX Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;

XX Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;

XX Salzberg SL;

```

XX DR WPI; 2002-557818/59.
XX
XX PT Novel DNA array useful for determining differential expression of
XX PT Methylococcus capsulatus genes, comprises polynucleotides or
XX PT oligonucleotides representative for a selective number of Methylococcus
XX PT capsulatus genes
XX
XX PS Claim 19; Page 81-82; 678pp; English.
XX
XX CC The invention relates to a novel DNA array giving a representation of a
XX CC number of Methylococcus capsulatus genes. The method of the invention is
XX CC useful for determination of the differential expression of the genes of
XX CC M. capsulatus, and for studying gene expression on a genomic scale and in
XX CC gene expression assays of M. capsulatus genes. The sequences shown in
XX CC ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the
XX CC invention.
XX
XX SQ Sequence 1098 BP; 156 A; 390 C; 374 G; 178 T; 0 other;
XX
XX Query Match 3.5%; Score 40.8; DB 24; Length 1098;
XX Best Local Similarity 48.7%; Pred. No. 0.018;
XX Matches 111; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
XX
XX QY 274 CTCTGGCGCTGGCTGCCATATCCGGAAGCGTTCCTCCGTCGGGATAAATAATCGCGCAGTG 333
XX Db 1003 CCGGCGCGCTGGCCATATCCCGCGCAGGTCCCGATCGTCGAGCAGCGCGGATGG 944
XX
XX QY 334 CGCGGTCCTCAGACATCCCGCAGCGGTACAGCGTCCCTGTACATCTCTCTGAA 393
XX Db 943 CGGCGCGCAGGCGCGCATCCCGGTGGGAGCAGTCCATTTCTCGCGTCGCGCA 884
XX
XX QY 394 TGACATCAGGATCCCGCGCTCTCACTGCGGATACGGGCGCAGGACTGTACGCTT 453
XX Db 883 CGGCTTCGGATCCCGCGCGCGGAGGCCAGATGGGAGCGCGCGCGCGCT 824
XX
XX QY 454 CAGCCAGTACCATACCAACGCTTCATTTCCGAGGAGCATGACCA 501
XX Db 823 GCAGCAGTGGATGCCAGGCGCTTCGCGCAGGCGCGGATGACCA 776
XX
XX RESULT 11
XX AAZ45259/c
XX ID AAZ45259 standard; DNA; 14245 BP.
XX AC AAZ45259;
XX
XX DT 27-MAR-2000 (first entry)
XX
XX DE DNA encoding enzymes involved in exopolysaccharide biosynthesis.
XX
XX KW Exopolysaccharide; EPS; ESP enzyme; EPS1; EPS2; EPS3; EPS4; EPS5; EPS6;
XX KW EPS7; EPS8; EPS9; EPS10; EPS11; Lactobacillus helveticus strain LH59;
XX KW activated D-galactose pyranose; saccharide; beta-glycosyltransferase;
XX KW undecaprenyl-phosphate-glycosyl-1-phosphate-transferase;
XX KW alpha-glycosyltransferase; EPS polymerase; glycosyltransferase;
XX KW phosphofuranose; transporter; food; fermented milk product; yoghurt;
XX KW cheese; flavour stability; organoleptic property; ss.
XX
XX OS Lactobacillus helveticus.
XX
XX FH Key Location/Qualifiers
XX FT CDS 1052..1729
XX FT /*tag= a
XX FT /product= "EPS1"
XX FT /note= "encodes AAY54075"
XX FT 1733..2848
XX FT /*tag= b
XX FT /product= "EPS2"
XX FT /note= "encodes AAY54076"
XX FT 2851..3942
XX FT /*tag= c
XX FT /product= "EPS3"

```

```

FT CDS /note= "encodes AAY54077"
FT 3929..5083
FT /*tag= d
FT /product= "EPS4"
FT /note= "encodes AAY54078"
FT 5076..6095
FT /*tag= e
FT /product= "EPS5"
FT /note= "encodes AAY54079"
FT 6098..7090
FT /*tag= f
FT /product= "EPS6"
FT /note= "encodes AAY54080"
FT 7095..8258
FT /*tag= g
FT /product= "EPS7"
FT /note= "encodes AAY54081"
FT 8283..9122
FT /*tag= h
FT /product= "EPS8"
FT /note= "encodes AAY54082"
FT 9135..10253
FT /*tag= i
FT /product= "EPS9"
FT /note= "encodes AAY54083"
FT 10250..11662
FT /*tag= j
FT /product= "EPS10"
FT /note= "encodes AAY54084"
FT 11664..13181
FT /*tag= k
FT /product= "EPS11"
FT /note= "encodes AAY54085; contains 1 stop codon at
FT nucleotides 12867-12869"
FT
FT WO9962316-A2.
FT
FT 09-DEC-1999.
FT
FT 22-APR-1999; 99WO-EP02841.
FT
FT 22-APR-1998; 98EP-0201310.
FT 22-APR-1998; 98EP-0201311.
FT 22-APR-1998; 98EP-0201312.
FT
FT (NEST ) SOC PROD NESTLE SA.
FT
FT Stingele F, Germond JE, Lamothe G;
FT
FT WPI; 2000-097267/08.
FT P-PSDB; AAY54075; AAY54076, AAY54077, AAY54078, AAY54079, AAY54080,
FT AAY54081; AAY54082, AAY54083, AAY54084; AAY54085.
FT
FT New recombinant enzymes for synthesis of exopolysaccharides,
FT particularly in lactic acid bacteria, for improving properties of
FT fermented milk products
FT
FT Claim 11; Page 110-117; 162pp; French.
FT
FT The present sequence encodes enzymes involved in the biosynthesis of
FT exopolysaccharides (EPS). These enzymes are designated EPS1-EPS11, and
FT are encoded by open reading frames eps1-eps11. The enzymes are isolated
FT from Lactobacillus helveticus strain LH59. The proteins are used
FT in a method for the synthesis of EPS, which includes at least one step
FT of forming a bond (alpha or beta-isomer) between C-1 (carrying the
FT reducing aldehyde function, of an activated D-galactose pyranose), and
FT a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis
FT of EPS occurs with, in each step, addition of a new sugar unit, through
FT its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar
FT unit, present at the end of a chain of sugar residues bonded to the
FT primer. EPS1 has homology with an undecaprenyl-phosphate-glycosyl-1-
FT phosphate-transferase; EPS2, EPS3 and EPS4 are alpha-
FT glycosyltransferases; EPS5 and EPS6 have homology with a beta-

```

CC	glycosyltransferase; EPS7 has homology with an EPS polymerase; EPS8
CC	is a glycosyltransferase; EPS9 catalyses the transfer of phosphofuranose
CC	onto the following repetitive unit; EPS10 transports the polysaccharides
CC	formed; and EPS11 is involved in the synthesis and export of formed
CC	polysaccharides. The EPS enzyme are used to improve properties of foods,
CC	particularly fermented milk products such as yoghurt and cheese,
CC	e.g. their organoleptic properties and flavour stability.
XX	
SQ	Sequence 14245 BP; 5095 A; 1755 C; 2545 G; 4850 T; 0 other;
Query Match 3.4%; Score 40.6; DB 21; Length 14245;	
Best Local Similarity 46.3%; Pred. No. 0.093;	
Matches 133; Conservative 0; Mismatches 154; Indels 0; Gaps 0;	
QY	226 CAATGATCTGCGGTTCTGTTAATCAACCGGTACTCAATATCTTCTCGCGCTGG 285
DB	2822 CACTATTTGTGGGACTCTTTTACCATCAAACTATTTTTTACAAATATTACTTTCAGTTT 2763
QY	286 CTGCGATCATCCGGAAGCGTTCGGTGGGGATAAAAATCGCGAGTGCGCCGTCATG 345
DB	2762 TAACAAAGAAGATCTCAAAATCAGGTGTGTTAAATAGTTTATCGATGCACTGCTGCTT 2703
QY	346 CAGACACATCCCCACGGGTAACAGGTCCTCTGCATTTCTTCTGAATGACATCAGGGA 405
DB	2702 TATGGTAATCATCTTTTATCAACTAGCATGCGATTAAACACCATTCGTAACGATATTAGGAA 2643
QY	406 TCCGCGCGCTCTCACTGGCGATTAACGGGACGCGGAGCTGACGCTTCAGCCAGTACCA 465
DB	2642 TAGCATCCCATTTAGTAGCTTACAATAGGCACTCCACAGCCATATACTCTGGTAGACTA 2583
QY	466 TACCAACGGTTCATTTTCCGAAGGCATGACCAACCACTGGCAATC 512
DB	2582 AACCACAGCCTCCCAACGAGATAGTACGATCCCCACATCCCATTAATC 2536

RESULT 12	
AAZ30356/c	
ID	AAZ30356 standard; DNA; 14245 BP.
XX	
XX	
AC	AAZ30356;
XX	
DT	11-FEB-2000 (first entry)
XX	
DE	Nucleotide sequence of the eps operon of L. helveticus LH59.
XX	
KW	eps operon; Lactobacillus helveticus LH59; enzyme; eps1; eps2;
KW	eps3; eps4; eps5; eps7; eps8; eps9; eps10; eps11;
KW	xopolysaccharide biosynthesis; EPS; intersugar bond; antitumour;
KW	probiotic; foodstuff; organoleptic quality; flavour;
KW	lactic acid bacteria; acidified milk product; yoghurt; cheese; ss.
XX	
OS	Lactobacillus helveticus.
XX	
Key	Location/Qualifiers
CDS	1052..1729
FT	/*tag= a
FT	/product= "eps1"
FT	/note= "encodes AAY43777; this CDS is specifically
FT	claimed (claim 11)."
CDS	1733..2848
FT	/*tag= b
FT	/product= "eps2"
FT	/note= "encodes AAY43778; this CDS is specifically
FT	claimed (claim 11)."
CDS	2851..3942
FT	/*tag= c
FT	/product= "eps3"
FT	/note= "encodes AAY43779; this CDS is specifically
FT	claimed (claim 11)."
CDS	3929..5083
FT	/*tag= d
FT	/product= "eps4"
FT	/note= "encodes AAY43780; this CDS is specifically
FT	

FT	CDS	5076..6095	claimed (claim 11). /tag= e /product= 'eps5' /note= 'encodes AAY43781; this CDS is specifically claimed (claim 11)'
FT	CDS	6098..7090	claimed (claim 11). /tag= f /product= 'eps6' /note= 'encodes AAY43782; this CDS is specifically claimed (claim 11)'
FT	CDS	7095..8258	claimed (claim 11). /tag= g /product= 'eps7' /note= 'encodes AAY43783; this CDS is specifically claimed (claim 11)'
FT	CDS	8283..9122	claimed (claim 11). /tag= h /product= 'eps8' /note= 'encodes AAY43784; this CDS is specifically claimed (claim 11)'
FT	CDS	9135..10253	claimed (claim 11). /tag= i /product= 'eps9' /note= 'encodes AAY43785; this CDS is specifically claimed (claim 11)'
FT	CDS	10250..11662	claimed (claim 11). /tag= j /product= 'eps10' /note= 'encodes AAY43786; this CDS is specifically claimed (claim 11)'
FT	CDS	11664..13181	claimed (claim 11). /tag= k /product= 'eps11' /note= 'encodes AAY43787; the sequence contains 1 stop codon at nucleotides 12867-12869; this CDS is specifically claimed (claim 11)'
XX	WO954475-A2.		
XX	28-OCT-1999.		
XX	22-APR-1999;	99WO-EP03011.	
XX	22-APR-1998;	98EP-0201310.	
PR	22-APR-1998;	98EP-0201311.	
PR	22-APR-1998;	98EP-0201312.	
XX			
PA	(NEST) SOC PROD NESTLE SA.		
XX	Stingele F, Germond JE, Lamothe G;		
PI	WPI; 2000-013255/01.		
DR	P-PSDB; AAY43777, AAY43778, AAY43779, AAY43780, AAY43781, AAY43782,		
DR	AAY43783, AAY43784, AAY43785, AAY43786, AAY43787.		
XX	New recombinant enzymes for biosynthesis of exopolysaccharides having e.g. antitumor or probiotic properties or useful in fermented milk products		
XX	Claim 11; Page 111-118; 163pp; French.		
XX	The present sequence represents the eps operon of Lactobacillus helveticus LH58. The operon contains 11 open reading frames, and encodes enzymes (eps1. eps2, eps3, eps4, eps5, eps6, eps7, eps8, eps9 and eps11) that are involved in the biosynthesis of exopolysaccharides (EPS). The enzymes catalyse the formation of specific intersugar bonds. The enzymes catalyse a process that includes at least one step of forming a bond (in alpha or beta anomeric form) between C1, carrying the reducing aldehyde group of an activated D-galp (galactose in pyranose form), and a phosphate group on a lipophilic or proteinaceous primer. The enzymes are used to produce EPS that have antitumor or probiotic properties or are used in foodstuffs to improve organoleptic properties.		

CC	nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and H37Rv (AAR99682). The method is useful for evaluating strain variation of M. tuberculosis and has valuable application in the fields of tuberculosiis genetics, epidemiology, patient treatment and epidemic monitoring.
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docId=6294328B1.
XX	
SQ	Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
	Query Match 3.4%; Score 39.8; DB 22; Length 4403765; Best Local Similarity 50.8%; Pred. No. 4.8; Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY	333 GCGCCGTTCATCGACAGACATCCCCCACCGGTAACACGCTCCCTGTGCACATTCTTCA 392
Db	577989 GCGTCGCCCACTGACCGACTCTGTCCCAGCACACCGGTGTCGGTGATCCGCTCGGC 577930
QY	393 ATGACATCAGGGATCCGCCCGTCTACTGGCGATAACGGGCACGCCGAGACTACGCT 452
Db	577929 ACCCGACGGCAGCCGCCACCGCCGCGGCCACACCGGTGTGCCGACGCTTGGGCC 577870
QY	453 TCAGCCAGTAGTACCATACCAAAGCTTCATTTTCCGAAGGCATGACCAACACACTGCGCAATC 512
Db	577869 TCCACAGCAACCAAGCCGCGAAGACTCGGAGTAGTCTCGCAGCCGCAACCAAGTCGCCGCC 577810
QY	513 CGGTAGA 519
Db	577809 CGAAACA 577803

RESULT 15
AAI99682/c
ID AAI99682 standard; DNA; 4411529 BP.
XX AC
XX AC
XX AC
XX AAI99682;
DT 15-JAN-2002 (first entry)
XX
XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
DE
XX
XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
XX Mycobacterium tuberculosis.
XX OS
XX
XX US6294328-B1.
PN
XX
XX 25-SEP-2001.
PD
XX
XX 24-JUN-1998; 98US-0103840.
PF
XX
XX 24-JUN-1998; 98US-0103840.
PR
XX
XX (GENO-) INST GENOMIC RES.
PA
XX
XX Fleischmann RD, White OR, Fraser CM, Venter JC;
PI
XX
XX WPI; 2001-647261/74.
DR
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ -
PT
XX
XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.

The invention relates to evaluating strain variation within and between different populations of the tuberculosis bacterial pathogen, *Mycobacterium tuberculosis* or related *Mycobacterium* by determining the nucleotide sequence of the first strain at positions in the complete sequence of the genome that correspond to positions that differ in the

nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and H37Rv (AAI99682). The method is useful for evaluating strain variation of M. tuberculosis and has valuable application in the fields of tuberculosis genetics, epidemiology, patient treatment and epidemic monitoring.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6294328B1.

Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

	Query Match	Score	% Identity	DB ID	Length	Mismatches	Gaps	Indels
QY	333 CGCGCCGCCATGACAGACATCCCCAACGGGTACACGCCTTCCTTCTGA	3.4%	50.8%	Pred. No. 4.8;	392	0	0	0
Db	576547 CGTGCGGCCACTGACGACCCTGTGCCCGCACACAGGGTCCCGGTGATCCC							
QY	393 ATGACATCAGGAGTCCCGCCGTTCACTGGCGATAACGGGACGCGAGACTGACGCT				452			
Db	576487 ACCGGAGGGCAGCGCCACCGCCGCGCCACCGGTGTCCGACGCTTGGGCC							
QY	453 TCAGCGAGTAGCATACCAAACGCTTCATTTCGAAGCATGACACACACTGGCAATC				512			
Db	576427 TCCACAGCAACACGGCCGACGAGTGGGAGTAGCTCGGCACCGAACCAGGTCCGCGCC							
QY	513 CGGTAGA 519							
Db	576367 CGAACA 576361							

RESULT 16
AAH41227

ID AAH41227 standard; DNA; 265118 BP.
XX AC
XX AC
XX AC
XX AC
XX OCT-2001 (first entry)
DE Pyrococcus abyssii genomic fragment #6.
XX KW
XX Hyperthermophilic archaeon; hyperthermophilic protein; ds.
OS Pyrococcus abyssi.
FH Key Location/Qualifiers
FT misc_feature 1..49980 /tag=a
FT FT /note="This sequence overlaps with the 3' end of
FT FT AAH41226"
XX PN FR2792651-Al.
XX XX
XX PD 27-OCT-2000.
XX PF 21-APR-1999; 99FR-0005034.
XX PR 21-APR-1999; 99FR-0005034.
XX PA (CNRS) CNRS CENT NAT RECH SCI.
XX PA (IFRE-) IFREMER INST FR RECH EXPL MER.
XX PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querrelou J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
XX PT New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX PS Claim 1; Page 593-665; 1657pp; French.
XX

CC The present invention relates to the genomic sequence of *Pyrococcus*
CC abyssus and *P. abyssus* proteins (see AAB96053-AAB96842). *P. abyssus* is a
CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal
CC vents. The present sequence is a fragment of the genomic sequence of *P.*
CC abyssus. The 5' end of this sequence overlaps with the 3' end of AAB41226.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAB75903-AAB75920 and AAG66436.
XX
SQ Sequence 265118 BP: 75226 A; 61302 C; 54005 G; 74585 T; 0 other;

Query Match 3.4%; Score 39.6; DB 22; Length 265118;
Best Local Similarity 52.4%; Pred. No. 1.1;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 344 TGCAGACATPCCCGGCGGTACAGCGTCCCTGTGCACATCTTCTGAATGACATCAGG 403
DB 57111 TTTAGCGCTCATCAATGCTATCAACAAGAACCGGTTCTTCCATCCCTAATCAACTCAGG 57170
QY 404 GATCCCGCGCTCTCACTGGGTATACGGCGGACCGGACACTGACGCTTCAGCCAGTAC 463
DB 57171 TATCCCGACGACTTCTCTTCCCAAGGCTGGGACAGAGAGGTTCCGCTCAAGGATTAC 57230
QY 464 CATACCAAGCGCTTCATTTCCGAAGCATGACACACACATCGCA 509
DB 57231 CATGCCAAGCGTTCCTCTCTAGAAAGCTAGAGACTAGGGCAAGTGCA 57276

RESULT 17
AAX84332
ID AAX84332 standard; DNA; 5059 BP.
XX
AC AAX84332;
XX
DT 08-SEP-1999 (first entry)
XX
DE Stealth virus nucleic acid clone, SEQ ID NO: 24.
XX
KW Stealth virus; detection; diagnosis; infection; ss.
XX
OS Stealth virus.
XX
FH Key Location/Qualifiers
FT misc_difference 3605
FT /tag= a
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3610
FT /tag= b
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3615
FT /tag= c
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3630
FT /tag= d
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3631
FT /tag= e
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3638
FT /tag= f
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"

FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3641
FT /tag= g
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3642
FT /tag= h
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3657
FT /tag= i
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3659
FT /tag= j
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3681
FT /tag= k
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3698
FT /tag= l
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3703
FT /tag= m
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3708
FT /tag= n
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3713
FT /tag= o
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3720
FT /tag= p
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3732
FT /tag= q
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3733
FT /tag= r
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3740
FT /tag= s
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3747
FT /tag= t
FT /note= "this nucleotide is represented as a * in the
FT specification, and is included to maintain the
FT base numbering given in the specification"


```

FT misc_difference 3750 /tag= u "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3751 /tag= v "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3766 /tag= w "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3778 /tag= x "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3792 /tag= y "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3799 /tag= z "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3800 /tag= aa "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3812 /tag= ab "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3813 /tag= ac "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3818 /tag= ad "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3819 /tag= ae "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3824 /tag= af "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3827 /tag= ag "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3841 /tag= ah "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3853 /tag= ai "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"

```

```

FT misc_difference 3857 /tag= aj "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3860 /tag= ak "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT
FT misc_difference 3863 /tag= al "this nucleotide is represented as a * in the
FT /note= specification, and is included to maintain the
FT base numbering given in the specification"
FT

Query Match 3.3%; Score 39.2; DB 20; Length 5059;
Best Local Similarity 14.5%; Pred. No. 0.15;
Matches 48; Conservative 138; Mismatches 146; Indels 0; Gaps 0;

QY 809 GGCCAGTCGAACCCGGTGAATACCGTTACCGGTGCTGCTGACACCTTCGCCCATCAG 868
Db 1564 SSSHWNTATBHCHVMTNSWHCHMATCHDVTMRTHSTATHVRSCNSARNDCATBYND 1623
QY 869 ATCCGCCATCATGGTGCAGATACGCACACAATCAATCAATCAATCAATCAATCAATCA 928
Db 1624 RNNGTMACTCHNGRGNSTHCHMVGNMRWHCHNCTDANDAMNCDSCNCDATAARVAABRR 1683
QY 929 CGTTCGGTCTTACGGGTGATGTAGTGTATTTTCTGTCGACAAATAGTGAAGCGGTGACAGCA 988
Db 1684 HSSCYMTGAVRSRHCVMWRACHCMARDTTHCRSNDNGSTATHVRSSNCTHSANAYSSSH 1743
QY 989 TATCACAGCGCTCAGTCTCTCTATATCTATCTATCTATCTATCTATCTATCTATCTATCT 1048
Db 1744 WDSAGNCANTYHGRHBASTNANDBASTHMGTYTHSTATHVRSSNCSTRHCVCMARDTHCM 1803
QY 1049 AGGTTAAATTCGCCGATATATCGTGAAGTCTGAGATGGAAGGAGGTGAAGGCTGT 1108
Db 1804 VRSWTHATYCHNCSNCSANADDTNACMARNSCDBMADWTHHGNBANKNTRSRSCMVNTH 1863
QY 1109 CCTGAAGGAATAAAAGTGCATCATGCCCTC 1140
Db 1864 SCASSANYNGRATRHMGYWASSNWTSCMVTH 1895

RESULT 18
AAH41226/c
ID AAH41226 standard; DNA: 349980 BP.
AC AAH41226;
XX
XX 29-OCT-2001 (first entry)
DE Pyrococcus abyssi genomic fragment #5.
XX Hyperthermophilic archaeon; hyperthermophilic protein; ds.
XX Pyrococcus abyssi.
OS
XX Key Location/Qualifiers
FH misc_feature 1..49980
FT /tag= a
FT /note= "this sequence overlaps with the 3' end of
FT AAH41225"
FT misc_feature 300001..349980
FT /tag= b
FT /note= "this sequence overlaps with the 5' end of
FT AAH41227"
XX
XX FR2792651-A1.
PN
XX 27-OCT-2000.
XX
XX 21-APR-1999; 99FR-0005034.
PF

```

XX 21-APR-1999; 99FR-0005034.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX (IFRE-) IFREMER INST FR RECH EXPL MER.
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
 XX Querellou J, Weissenbach J, Saurin W, Heilig R;
 XX WPI; 2001-126236/14.
 XX New nucleotide sequences isolated from *Pyrococcus abyssi* encode
 XX proteins useful in industry -
 XX Claim 1; Page 511-606; 1657pp; French.
 XX The present invention relates to the genomic sequence of *Pyrococcus*
 XX *abyssi* and *P. abyssi* proteins (see AAB96053-AAB96842). *P. abyssi* is a
 XX hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal
 XX vents. The present sequence is a fragment of the genomic sequence of *P.*
 XX *abyssi*. The 5' end of this sequence overlaps with the 3' end of AAB41225
 XX and the 3' end of this sequence overlaps with the 5' end of AAB41227. The
 XX proteins of the present invention have various potential industrial uses,
 XX since the proteins are stable at very high temperatures, some up to 110
 XX degrees centigrade.
 XX Note: This patent is in the same patent family as WO200065062, which
 XX contains additional sequences as shown in AAB99132-AAB99143.
 XX AAB75903-AAH75920 and AAG66436.
 XX Sequence 349980 BP; 98084 A; 80447 C; 77665 G; 93784 T; 0 other;
 SQ
 Query Match 3.3%; Score 39.2; DB 22; Length 349980;
 Best Local Similarity 54.1%; Pred. No. 1.7;
 Matches 80; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 333 GCGCGCGTCCATGACACATCCCGGTAACAGCGTCCCTGTTCACATCTTCTGA 392
 DB 224168 GCTTCGCGCAATCTAAATGGATCTTTAGTGGAAACAGTATCCCGAGTCTTCCATCTCG 224109
 QY 393 ATGACATCAGGATCCCGCGTCTCACTGGCGATACGGGACGCGGAGACTGACGCT 452
 DB 224108 ATGACATGCTTTATACCTCCGACATTTGTTCCATAACTGCTGCCACTAGCACCGCT 224049
 QY 453 TCAGCGAGTACCATACCAACGCTTCAT 480
 DB 224048 TCTATCAGAACCATTTCCAAACCCCTCTT 224021

RESULT 19
 ABQ90189/c
 ID ABQ90189 standard; DNA; 1110 BP.
 AC ABQ90189;
 DT 01-OCT-2002 (first entry)
 DE M. capsulatus gene #174 for DNA array.
 DE Micro array; gene; ds; differential expression; gene expression.
 KW Methyllococcus capsulatus.
 OS Methyllococcus capsulatus.
 XX WO200255655-A2.
 PN 18-JUL-2002.
 XX 14-JAN-2002; 2002WO-NO00019.
 XX 12-JAN-2001; 2001NO-0000235.
 XX 12-JAN-2001; 2001NO-0000239.
 XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
 PA (TIGR-) TIGR.

XX Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
 PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
 PI Salzberg SL;
 XX WPI; 2002-557818/59.
 XX Novel DNA array useful for determining differential expression of
 PT Methyllococcus capsulatus genes, comprises polynucleotides or
 PT oligonucleotides representative for a selective number of Methyllococcus
 PT capsulatus genes -
 XX Claim 19; Page 135-136; 678pp; English.
 XX The invention relates to a novel DNA array giving a representation of a
 CC number of Methyllococcus capsulatus genes. The method of the invention is
 CC useful for determination of the differential expression of the genes of
 CC M. capsulatus, and for studying gene expression on a genomic scale and in
 CC gene expression assays of M. capsulatus genes. The sequences shown in
 CC ABO90016-ABO91855 represent M. capsulatus genes for use in arrays of the
 CC invention.
 XX Sequence 1110 BP; 160 A; 344 C; 403 G; 203 T; 0 other;
 SQ
 Query Match 3.3%; Score 38.4; DB 24; Length 1110;
 Best Local Similarity 50.5%; Pred. No. 0.11;
 Matches 93; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
 QY 325 CGCGCAGTGGCGGTCATGACACATCCCGGTAACAGCGTCCCTGTTCACAT 384
 DB 958 CGACACGAGTCCGCGGAGTCTTGTGACGTCGCGCGGCGGACGACGCGCATTC 899
 QY 385 TCTTCTGAATGACATCAGGATCCCGCGTCTCACTGGCGATACGGGACGCGGAGA 444
 DB 898 CCTCGCGAGGAGTTCGGGTATGCCCGGACCGGCTGCGCCAGCAGCGGACCG 839
 QY 445 CTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAGGATGACCCACAC 504
 DB 838 CCATCGCTTCCAGATGACAGCGGCGGAGCCCTTCGAGTAGGAGGAGGCGGCGCC 779
 QY 505 TGGC 508
 DB 778 GGGC 775
 RESULT 20
 ABK74439/c
 ID ABK74439 standard; DNA; 1108 BP.
 AC ABK74439;
 XX 13-AUG-2002 (first entry)
 DE Bacillus licheniformis genomic sequence tag (GST) #1730.
 DE Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX Bacillus licheniformis.
 XX WO200229113-A2.
 XX 11-APR-2002.
 XX 05-OCT-2001; 2001WO-US31437.
 XX 06-OCT-2000; 2000US-0680598.
 XX 27-MAR-2001; 2001US-279526P.
 XX (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX

PI Berka R, Clausen IG;
DR WPI; 2002-416684/44.
XX
XX
XX Monitoring differential expression of several genes in first Bacillus
XX cell relative to expression of same genes in one or more second
XX Bacillus cells, by using substrate containing Bacillus genomic
XX sequenced tag array -
PS Claim 4; SEQ ID NO 1730; 200pp; English.
XX
XX The invention describes a method of monitoring differential expression of
XX genes in a first Bacillus cell relative to expression of the genes in
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes
XX isolated from Bacillus cells to a substrate containing array of Bacillus
XX genomic sequenced tags (GST), examining the array, and determining
XX relative gene expression by an observed hybridisation reporter signal of
XX a spot in the array. The method is useful for measuring the expression of
XX genes in a first Bacillus cell relative to expression of the same genes
XX in one or more second Bacillus cells. The method is useful for monitoring
XX global expression of several genes from a Bacillus cell, discovering new
XX genes, identifying possible functions of unknown open reading frames and
XX monitoring gene copy number variation and stability. Monitoring changes
XX in expression of genes may be used to provide a representation of the way
XX in which Bacillus cells adapt to changes in culture conditions, as
XX environmental stress or other physiological provocation. Extensive
XX follow-up characterisation is unnecessary, when one spot on an array
XX equals one gene or one open reading frame, since sequence information is
XX available. This sequence represents a genomic sequence tag (GST) used in
XX the method of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1108 BP; 304 A; 290 C; 246 G; 268 T; 0 other;

Query Match 3.1%; Score 37; DB 24; Length 1108;
Best Local Similarity 54.9%; Pred. No. 0.33;
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 387 TTCTGAATGACATCAGGGATCCGCCGCTCTCACTGGCGATACGCGGAGACT 446
DB 926 TTCTGTATTAATCTTTCAGGCGCGGAGTCTGCCCGATACAGCGCTCCGCGAGCG 867
QY 447 GACGCTTCCAGGAGTACCATACAAACGCTTTCATTTCCGAGGCGATGACACACATG 506
DB 866 AGCGCTTCAAGGCGGAATTCGAGAGTTTCGTCGGAGAGGAAGACGAACAATCT 807
QY 507 GCATCCGCTAGA 519
DB 806 GAGCTCGCATAGA 794

RESULT 21
ABV47264
ID ABV47264 standard; cDNA; 386 BP.
XX
XX ABV47264;

16-SEP-2002 (first entry)

Human prostate expression marker cDNA 47255.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

XX

PF 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR

XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 9300; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX
XX Sequence 386 BP; 110 A; 97 C; 93 G; 85 T; 1 other;
Query Match 3.1%; Score 36.6; DB 23; Length 386;
Best Local Similarity 56.1%; Pred. No. 0.24;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1 CTGCAGGAGATGAAAAAGCCAAATAAAATTTGCCATCCCGAGGCTCCAGCTG 60
DB 257 CTGAAGCAAAAAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 316
QY 61 AAAGTAGGCTGTCTCTCGGTATTTAAATGATTCGTCCTCCGCTATTTAAACAATG 120
DB 317 CAATTCGCCCTATAGTAGTCTGATTCAGCGGCTCACTGGCGCTGTTTACACGCTG 376
QY 121 TGA 123
DB 377 TGA 379

RESULT 22
AAD22921
ID AAD22921 standard; cDNA; 654 BP.
XX
XX AAD22921;

26-FEB-2002 (first entry)

Human soluble Zalphall cytokine receptor cDNA degenerate sequence.

Human; Zalphall; cytokine receptor; immunosuppressive; cytostatic;
Inflammatory disorder; haemostatic; cell proliferation; immune disorder;
autoimmune disease; rheumatoid arthritis; multiple sclerosis; cancer;
myasthenia gravis; systemic lupus erythematosus; SLE; diabetes; asthma;
ulcerative colitis; inflammatory bowel disease; Crohn's disease; sepsis;
viral infection; ss.

XX

DT 17-OCT-2001 (first entry)
 XX Probe #8926 used to measure gene expression in human placenta sample.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX Homo sapiens.
 OS
 XX WO200157272-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00663.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 PT Claim 25; SEQ ID No 8926; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX Sequence 598 BP; 169 A; 109 C; 137 G; 183 T; 0 other;
 SQ

Query Match 3.1%; Score 36.2; DB 22; Length 598;
 Best Local Similarity 56.2%; Pred. No. 0.42;
 Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1053 TTAATTCGCCGATAATCCGTCGAAGTCTGAGGATGAAGGAGGCTGTTCTG 1112
 DB 503 TAAATGTACAGGCCACTTGTACAGAGCTCATGATGACAGTGAATGAACACTCATGTC 444
 QY 1113 AAAGGAATAAAGTGACATCATGCGCTCTTTTCTGCTCCGAGGCAATTTTACTTTT 1172
 DB 443 AACAGATAAAGTGACCTAATACATCTCTGACAGGTTTCTTGAGTTTGGATATTTT 384
 QY 1173 T 1173
 DB 383 T 383

RESULT 27
 AAD22970
 ID AAD22970 standard; cDNA; 741 BP.
 XX
 AC AAD22970;
 XX
 XX 26-FEB-2002 (first entry)
 DT Human soluble zalphall cytokine receptor variant degenerate cDNA.
 XX Human; Zalphall; cytokine receptor; immunosuppressive; cytostatic;
 KW inflammatory disorder; haemostatic; cell proliferation; immune disorder;
 KW autoimmune disease; rheumatoid arthritis; multiple sclerosis; cancer;
 KW myasthenia gravis; systemic lupus erythematosus; SLE; diabetes; asthma;

KW ulcerative colitis; inflammatory bowel disease; Crohn's disease; sepsis;
 XX viral infection; variant; ss.
 OS Homo sapiens.
 XX WO200177171-A2.
 PN 18-OCT-2001.
 XX 03-APR-2001; 2001WO-US10872.
 XX 05-APR-2000; 2000US-194731P.
 PR 28-JUL-2000; 2000US-222121P.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Sprecher CA, Novak JE, West JW, Presnell SR, Holly RD, Nelson AJ;
 PI WPI; 2002-025898/03.
 XX Novel soluble receptor polypeptides and polynucleotides used as
 PT cytokine antagonist for stimulating ligand activity-induced
 PT proliferation of hematopoietic cells and for suppressing immune
 PT response in a mammal
 XX Disclosure; Page 222-223; 243pp; English.
 XX The invention relates to an isolated soluble zalphall cytokine receptor
 CC polypeptide and their cDNA molecules. Zalphall proteins are useful for
 CC inhibiting or antagonising the ligand activity-induced proliferation of
 CC hematopoietic cells and haematopoietic cell progenitors preferably
 CC lymphoid cells which are natural killer cells or cytotoxic T cells.
 CC Zalphall is useful for treating immune and inflammatory disorders, for
 CC reducing proliferation of neoplastic B or T cells, for suppressing an
 CC immune response in a mammal exposed to an antigen or pathogen. Zalphall is
 CC useful for treating diseases that require immune regulation including
 CC autoimmune diseases such as rheumatoid arthritis, multiple sclerosis,
 CC myasthenia gravis, systemic lupus erythematosus (SLE) and diabetes;
 CC asthma, ulcerative colitis, inflammatory bowel disease, Crohn's disease,
 CC sepsis, viral infection (dengue virus infection) and cancer. The present
 CC sequence is human soluble zalphall cytokine receptor variant cDNA
 CC degenerate sequence.
 XX Sequence 741 BP; 133 A; 79 C; 116 G; 91 T; 322 other;
 SQ

Query Match 3.1%; Score 36.2; DB 24; Length 741;
 Best Local Similarity 27.6%; Pred. No. 0.48;
 Matches 83; Conservative 48; Mismatches 170; Indels 0; Gaps 0;

QY 833 AACCGTTACGGGTGCTCTGACACCTTCCGCCATCAGATCGCCATCATGGTGAGATAGG 892
 DB 139 AAYTNCAYCCNWSNACNTNACNTNACNTGATHTNWSNAAAYACNGGNTGATY 198
 QY 893 CACAACAATGAATACACAGATAATTCAGGGAACAGTTCTGCTTACGGGTGATGA 952
 DB 199 ATHAARGAYMGACNTNGAYTTCAGGAGCARTAYGARGARYTNAARGAYGAGCN 258
 QY 953 GGTTTTTGTCTGACATACTAGTGAAGCGGTGACAGCATATCAGACGCTCAGTCTCTAT 1012
 DB 259 ACNWSNTGYWSNTNCAYMGNWSNCAAYAAACNACNCAACNTATACNTGYCAY 318
 QY 1013 ATTACTGTCTATGCCACTATGCGCAGATGACAGATGAGTTTAAATCCCGGATAATCCG 1072
 DB 319 ATGGAYGTNTTYTATGTCNGAYGATHTTYWSNCTNAAATACNACNACNWSN 378
 QY 1073 TCGAAGTCTGAGGATGAAGGAGGCTGCTTCTCCCTGAAAGGATATAAAGTACATC 1132
 DB 379 GGAAATAYNCAARGARTGCGNWSNTTYYTNTGNCNCAWSNMGNCARTAYATYATH 438
 QY 1133 A 1133
 DB 439 W 439

Db 379 ATGAGTNTTCATYTTATGGGNGAYGATHTTWSNCTNAAYATHACNGAYCARWSN 438
 QY 1073 TCGAAGTCTGAGTGAAGAAGGTTGAGGCTGTTCCCTGAAAGGAATAAAAGTGCATC 1132
 Db 439 CGNAAYTAYWNCARGARTGYGWNSTNTTYTNGCNGARWSNMGNCARTAYAYATH 498
 QY 1133 A 1133
 Db 499 W 499
 RESULT 30
 ID AAL10594/C
 AC AAL10594;
 DT 29-JUN-2000 (first entry)
 DE Gene encoding a subunit of cellulose synthase.
 XX Cellulose synthase; cellulose production; increase yield; ds.
 XX Vigna angularis.
 XX JP2000060568-A.
 XX 29-FEB-2000.
 XX 26-AUG-1998; 98JP-0239998.
 XX 26-AUG-1998; 98JP-0239998.
 XX (MIZU/) MIZUNO K.
 XX (OJIP) OJI PAPER CO.
 XX WPI; 2000-342371/30.
 XX P-PSDB; AY85179.
 XX A gene encoding a cellulose synthetic equipment - for the improvement
 PT in the amount of cellulose synthesised in a plant body
 XX Claim 2; Page 14-21; 32pp; Japanese.
 XX This sequence represents a gene encoding a subunit of the cellulose
 CC synthase complex of Vigna angularis. The invention relates to subunits of
 CC cellulose synthetic equipment, that can be used to increase the amount of
 CC cellulose synthesised by a plant. The proteins and genes encoding them
 CC can also be used to improve the properties of the cellulose being
 CC produced by a plant.
 XX Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;
 SQ
 Query Match 3.1%; Score 36.2; DB 21; Length 10732;
 Best local similarity 15.4%; Pred. No. 2.3;
 Matches 68; Conservative 174; Mismatches 197; Indels 3; Gaps 1;
 QY 178 TCGCGGTTATCTGTTCATTTCCCTGAGCTGACTAGCCAGACAGACAATGATCTGTG 237
 Db 10137 DNGSRSTSTCTYAKYST 10078
 QY 238 CGGTTCTGTAATATCAACCGGTACTCAATATCTCTCTGCGGCTGCGTGCATFATCC 297
 Db 10077 CYRASYDASRRCSYSTYASRGTSDDCCTBSRYSCYSYASRYANCDTBCYTTBARY- 10019
 QY 298 GGAAGCTTCGGTGGGATAAATAATCCGAGTCCGCCGCTCATCAGACACATCCC 357
 Db 10018 --RARCIDAYAKRCSSTSYRAYSSTCTYRCRRCNCSYSTSYSTSYSTSYSTSYST 9961
 QY 358 CCACGGGTACACGCGCTCTGTCATATCTCTGATGACATCAGGATCCCGCCGCT 417
 Db 9960 CTBCSRRCYSSRYSSFCNCYSYCCYTSYSTSTSTSTSTSTSTSTSTSTSTSTSTST 9901

QY 418 CACTGGCGATAACGGCGACGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTT 477
 Db 9900 SRGYSRSGDSRGNCYNTNCTDSTDBYSRCCYTYSTSTSTSTSTSTSTSTSTSTSTSD 9841
 QY 478 CATTTTCCGAAGCATGACCAACACACATCGGCAATCCGGTAGACCGGTACCGTGGGAAA 537
 Db 9840 CVTTTBSRSTSDSTSTYRCSRSDYDAYBSDNSTNCCYDASRTBTBSTNCACTBYDA 9781
 QY 538 GGGCACCTGCCATTAACACATCTCGCTCATCCAGGCTGTTCTGTCTGCTGACGACAC 597
 Db 9780 RCSRSTSYSSRGYDANSTSRYSSTSYSSSTSYSSSTSYSSSTSYSSSTSYSSSTSYSS 9721
 QY 598 GTGCTTCGTATTCTTCACGCC 619
 Db 9720 YDANCYSSDSTYBYCSRCCC 9699
 RESULT 31
 ID ABV03347
 AC ABV03347;
 DT 13-SEP-2002 (first entry)
 DE Human prostate expression marker cdNA 3338.
 XX Human prostate expression marker cdNA 3338.
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 XX WO200160860-A2.
 XX 23-AUG-2001.
 XX 20-FEB-2001; 2001WO-US05171.
 XX 17-FEB-2000; 2000US-183319P.
 XX 16-MAR-2000; 2000US-189862P.
 XX 25-MAY-2000; 2000US-207454P.
 XX 09-JUN-2000; 2000US-211314P.
 XX 18-JUL-2000; 2000US-219007P.
 XX 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1; Page 609; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 650 BP; 180 A; 110 C; 118 G; 199 T; 43 other;
 Query Match 3.0%; Score 36; DB 23; Length 650;
 Best Local Similarity 49.7%; Pred. No. 0.51;
 Matches 75; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 14 AAAAAAGCCAAATTAATAATGCGCATCCAGCGCTCCAGCTGAAGTAGGCTGT 73
 DB 479 AAAAAAGGGGAAAAAATAATTAANAACCCCGGCTTTTAAAAAATNGNNAAT 538
 QY 74 TCTGTCCGGTATTAAATGATGACCGCTCCCGTATTAACAATGTGATAAATTAATC 133
 DB 539 NGTTATTAATAAATTTNTTCCCGCCCGCATTTTAACCAATNCCCTTTTCCAT 598
 QY 134 CGTTACGGAAACCGCTGAACAAAATTCGG 164
 DB 599 GGTAAATTTAAACCAATTAANAAGGGG 629
 RESULT 32
 ABK73466/c
 ID ABK73466 standard; DNA; 1427 BP.
 AC ABK73466;
 XX
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #757.
 KW
 KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 XX Bacillus licheniformis.
 XX WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-US31437.
 XX
 XX 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 PA (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 XX Berka R, Clausen IG;
 XX WPI; 2002-416684/44.
 DR
 XX Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 XX
 PS Claim 4; SEQ ID NO 757; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,

CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1427 BP; 378 A; 305 C; 379 G; 365 T; 0 other;
 Query Match 3.0%; Score 36; DB 24; Length 1427;
 Best Local Similarity 46.0%; Pred. No. 0.81;
 Matches 160; Conservative 0; Mismatches 185; Indels 3; Gaps 1;
 QY 398 ATCAGGATCCCGCGCTGTCTACATGCGGATAAGGGGACCGGAGACTGACGCTTCAGC 457
 DB 1215 ATCATGAAGCCCGCTGTTCCTGTACACCGGACCGGCGCTGAGGCGAATCAG 1156
 QY 458 CAGTACCATACCAACGCTTCATTTTCCGAAGCATGACCAACACACTGGCAATCCGGTA 517
 DB 1155 CTGTCCAAGACCGCACGGCTCGAATTTTGAAGCATCAGGAAAGATCGCTCGCGGCATA 1096
 QY 518 GACCGGTAACGCTGGGAAAGGCGACCTGCCATTAACACATCTCCGCTCATTCGCCAGTG 577
 DB 1095 AATCTTTCGGCTGAGCGGTTTCATGAACACGATATACGCTTGCACCTGTCCGGAAGC 1036
 QY 578 TTCTGTCTGTGACGAGAGCGTGTCTGATTTCTTCAGCGCGCGCCACACGAGCCA 637
 DB 1035 ATGCTCCATATACCGGAAGTAATTTTCAAAATTCGGTTCCTCCGTTTCTTAAACGATCAG 976
 QY 638 GCGAAATGATTTCCCTTCCATCTTCAGCTGATACATACACGAGCATTAATTCATGTC 697
 DB 975 CTGCATGTCTTCTTCTTCA---ACAGTTTCATCAATACCGCTGTATCAATCGAAGCC 919
 QY 698 TTTTTCGGGACGTAGCATCCCACTGACGATAAGCGGAACATTTGTC 745
 DB 918 TTTTTCGTCGTCAGCGCTTTCACCATGCTGTATGAGCGGAATGCTTTC 871
 RESULT 33
 AAH29949/c
 ID AAH29949 standard; DNA; 1859 BP.
 XX
 XX AAH29949;
 XX
 DT 27-JUL-2001 (first entry)
 XX
 DE S cerevisiae apoptosis associated coding sequence YKR040C.
 XX
 KW Yeast; fungus; apoptosis; infection; proliferative disease;
 KW vaccine; autoimmune disease; ischaemia; neurodegeneration; ds.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO200102550-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 03-JUL-2000; 2000WO-BE00077.
 XX
 PR 01-JUL-1999; 99EP-0870141.
 XX
 PA (JANC) JANSSEN PHARM NV.
 XX
 PI Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
 PI Nelissen BJM, Reekmans RJ;
 XX
 DR WPI; 2001-367042/38.
 DR P-PSDB; AAG70913.
 XX
 PT Yeast and fungal nucleic acids encoding proteins involved in a pathway

PT leading to programmed cell death, useful for treating proliferative
 PT disorders, yeast and fungal infections, or for preventing apoptosis in
 XX certain diseases
 PS Claim 1; Fig 1; 218pp; English.
 XX The present invention provides the protein and coding sequences of a
 CC number of apoptosis associated proteins from the yeast *Saccharomyces*
 CC *cerevisiae* and the fungus *Candida albicans*. These can be used to identify
 CC treatments for fungal and yeast infections, for proliferative diseases
 CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
 CC and neurodegeneration. The present sequence is one of the *S. cerevisiae*
 CC coding sequences of the invention.
 XX
 SQ Sequence 1859 BP; 663 A; 297 C; 358 G; 541 T; 0 other;

Query Match 3.0%; Score 36; DB 22; Length 1859;
 Best Local Similarity 56.9%; Pred. No. 0.95;
 Matches 66; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 375 CCTGTGACATCTCTGAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGC 434
 DB 1459 CTGTGATCTTCTTGGTAAGACTTCGGGTATGCTTCCCACTTTTGTAGTGACACATAT 1400
 QY 435 ACGCCGGAGACTGACGTTTACGAGTACCATACCAACGCTTCATTTCCGAAGG 490
 DB 1399 AACCCACATGATGACGCTTCCACAATACTGTACCAACGCTGTGTTAATGAAGG 1344

RESULT 34
 AAS65293/c
 ID AAS65293 standard; cDNA; 426 BP.
 XX
 AC AAS65293;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #1097.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR
 DR P-PSDB; ABG01106.
 XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID No 1097; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 426 BP; 113 A; 105 C; 113 G; 95 T; 0 other;
 Query Match 3.0%; Score 35.8; DB 23; Length 426;
 Best Local Similarity 65.8%; Pred. No. 0.47;
 Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 525 AACGCTGGGAAAGGCGACCTGCCATTACACATCTCCGCTCATCCAGGTGTCTGTC 584
 DB 394 AATCTTAGACAGACGACCTTCCAGGACACAGCTCCCCCCTACTGCTGTGACCTTC 335
 QY 585 TGCTGACGACGACGCTGCTT 603
 DB 334 CTCTGGGAAGACTGGCTT 316

RESULT 35
 ABV48872
 ID ABV48872 standard; cDNA; 488 BP.
 XX
 AC ABV48872;
 XX
 DT 17-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 48863.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX

PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 9567; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX DE Human nervous system related polynucleotide SEQ ID NO 11358.
XX DE
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW KW Immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX OS
XX PN WC200159063-A2.
XX PD
XX PD 16-AUG-2001.
XX PD
XX PD 17-JAN-2001; 2001WO-US01334.
XX PD
XX PD 31-JAN-2000; 2000US-0179065.
XX PD 04-FEB-2000; 2000US-0180828.
XX PD 24-FEB-2000; 2000US-0184664.
XX PD 02-MAR-2000; 2000US-0186350.
XX PD 16-MAR-2000; 2000US-0189874.
XX PD 17-MAR-2000; 2000US-0190076.
XX PD 18-APR-2000; 2000US-0198123.
XX PD 19-MAY-2000; 2000US-0205515.
XX PD 07-JUN-2000; 2000US-0209467.
XX PD 28-JUN-2000; 2000US-0214886.
XX PD 30-JUN-2000; 2000US-0215135.
XX PD 07-JUL-2000; 2000US-0216647.
XX PD 07-JUL-2000; 2000US-0216880.
XX PD 11-JUL-2000; 2000US-0217487.
XX PD 11-JUL-2000; 2000US-0217496.
XX PD 14-JUL-2000; 2000US-0218290.
XX PD 26-JUL-2000; 2000US-0220963.
XX PD 26-JUL-2000; 2000US-0220964.
XX PD 14-AUG-2000; 2000US-0224518.
XX PD 14-AUG-2000; 2000US-0224519.
XX PD 14-AUG-2000; 2000US-0225213.
XX PD 14-AUG-2000; 2000US-0225214.
XX PD 14-AUG-2000; 2000US-0225266.
XX PD 14-AUG-2000; 2000US-0225267.
XX PD 14-AUG-2000; 2000US-0225268.
XX PD 14-AUG-2000; 2000US-0225270.
XX PD 14-AUG-2000; 2000US-0225447.
XX PD 14-AUG-2000; 2000US-0225757.
XX PD 14-AUG-2000; 2000US-0225758.
XX PD 14-AUG-2000; 2000US-0225759.
XX PD 18-AUG-2000; 2000US-0226279.
XX PD 22-AUG-2000; 2000US-0226681.
XX PD 22-AUG-2000; 2000US-0226868.
XX PD 23-AUG-2000; 2000US-0227182.
XX PD 30-AUG-2000; 2000US-0228924.
XX PD 01-SEP-2000; 2000US-0229287.
XX PD 01-SEP-2000; 2000US-0229343.
XX PD 01-SEP-2000; 2000US-0229344.
XX PD 01-SEP-2000; 2000US-0229345.
XX PD 05-SEP-2000; 2000US-0229509.
XX PD 05-SEP-2000; 2000US-0229513.
XX PD 08-SEP-2000; 2000US-0230437.
XX PD 08-SEP-2000; 2000US-0230438.
XX PD 08-SEP-2000; 2000US-0231242.
XX PD 08-SEP-2000; 2000US-0231243.
XX PD 08-SEP-2000; 2000US-0231244.
XX PD 08-SEP-2000; 2000US-0231413.
XX PD 08-SEP-2000; 2000US-0231414.
XX PD 08-SEP-2000; 2000US-0232080.
XX PD 08-SEP-2000; 2000US-0232081.
XX PD 12-SEP-2000; 2000US-0231968.
XX PD 14-SEP-2000; 2000US-0232397.
XX PD 14-SEP-2000; 2000US-0232398.
XX PD 14-SEP-2000; 2000US-0232399.
XX PD 14-SEP-2000; 2000US-0232400.
XX PD 14-SEP-2000; 2000US-0232401.
XX PD 14-SEP-2000; 2000US-0233063.
XX PD 14-SEP-2000; 2000US-0233064.
XX PD 14-SEP-2000; 2000US-0233065.
XX PD 21-SEP-2000; 2000US-0234223.
XX PD 21-SEP-2000; 2000US-0234274.
XX PD 25-SEP-2000; 2000US-0234997.
XX PD 25-SEP-2000; 2000US-0234998.
XX PD 26-SEP-2000; 2000US-0235484.
XX PD 27-SEP-2000; 2000US-0235834.
XX PD 27-SEP-2000; 2000US-0235836.
XX PD 29-SEP-2000; 2000US-0236327.
XX PD 29-SEP-2000; 2000US-0236367.
XX PD 29-SEP-2000; 2000US-0236368.
XX PD 29-SEP-2000; 2000US-0236369.
XX PD 29-SEP-2000; 2000US-0236370.
XX PD 02-OCT-2000; 2000US-0236802.
XX PD 02-OCT-2000; 2000US-0237037.
XX PD 02-OCT-2000; 2000US-0237038.
XX PD 02-OCT-2000; 2000US-0237039.
XX PD 02-OCT-2000; 2000US-0237040.
XX PD 13-OCT-2000; 2000US-0239935.
XX PD 13-OCT-2000; 2000US-0239937.
XX PD 20-OCT-2000; 2000US-0240960.
XX PD 20-OCT-2000; 2000US-0241785.
XX PD 20-OCT-2000; 2000US-0241786.
XX PD 20-OCT-2000; 2000US-0241787.
XX PD 20-OCT-2000; 2000US-0241808.
XX PD 20-OCT-2000; 2000US-0241809.
XX PD 20-OCT-2000; 2000US-0241826.
XX PD 20-OCT-2000; 2000US-0242221.
XX PD 01-NOV-2000; 2000US-0244617.
XX PD 08-NOV-2000; 2000US-0246474.
XX PD 08-NOV-2000; 2000US-0246475.
XX PD 08-NOV-2000; 2000US-0246476.
XX PD 08-NOV-2000; 2000US-0246477.
XX PD 08-NOV-2000; 2000US-0246478.
XX PD 08-NOV-2000; 2000US-0246523.
XX PD 08-NOV-2000; 2000US-0246524.
XX PD 08-NOV-2000; 2000US-0246525.
XX PD 08-NOV-2000; 2000US-0246526.
XX PD 08-NOV-2000; 2000US-0246527.
XX PD 08-NOV-2000; 2000US-0246528.
XX PD 08-NOV-2000; 2000US-0246532.
XX PD 08-NOV-2000; 2000US-0246609.
XX PD 08-NOV-2000; 2000US-0246610.
XX PD 08-NOV-2000; 2000US-0246611.
XX PD 08-NOV-2000; 2000US-0246613.
XX PD 17-NOV-2000; 2000US-0249207.
XX PD 17-NOV-2000; 2000US-0249208.
XX PD 17-NOV-2000; 2000US-0249209.
XX PD 17-NOV-2000; 2000US-0249210.
XX PD 17-NOV-2000; 2000US-0249211.
XX PD 17-NOV-2000; 2000US-0249212.
XX PD 17-NOV-2000; 2000US-0249213.
XX PD 17-NOV-2000; 2000US-0249214.
XX PD 17-NOV-2000; 2000US-0249215.
XX PD 17-NOV-2000; 2000US-0249216.
XX PD 17-NOV-2000; 2000US-0249217.
XX PD 17-NOV-2000; 2000US-0249218.
XX PD 17-NOV-2000; 2000US-0249219.
XX PD 17-NOV-2000; 2000US-0249244.
XX PD 17-NOV-2000; 2000US-0249245.
XX PD 17-NOV-2000; 2000US-0249246.
XX PD 17-NOV-2000; 2000US-0249265.
XX PD 17-NOV-2000; 2000US-0249297.
XX PD 17-NOV-2000; 2000US-0249299.
XX PD 01-DEC-2000; 2000US-0249300.
XX PD 01-DEC-2000; 2000US-0250391.
XX PD 05-DEC-2000; 2000US-0251160.
XX PD 05-DEC-2000; 2000US-0251030.
XX PD 05-DEC-2000; 2000US-0251988.

PD	XX	09-AUG-2001.
PF	XX	17-JAN-2001; 2001WO-US01354.
XX	XX	31-JAN-2000; 2000US-0179065.
PR	XX	04-FEB-2000; 2000US-0180628.
PR	XX	24-FEB-2000; 2000US-0184654.
PR	XX	02-MAR-2000; 2000US-0186350.
PR	XX	16-MAR-2000; 2000US-0189874.
PR	XX	17-MAR-2000; 2000US-0190076.
PR	XX	18-APR-2000; 2000US-0198123.
PR	XX	19-MAY-2000; 2000US-0205515.
PR	XX	07-JUN-2000; 2000US-0209467.
PR	XX	28-JUN-2000; 2000US-0214886.
PR	XX	30-JUN-2000; 2000US-0215135.
PR	XX	07-JUL-2000; 2000US-0216647.
PR	XX	07-JUL-2000; 2000US-0216880.
PR	XX	11-JUL-2000; 2000US-0217487.
PR	XX	11-JUL-2000; 2000US-0217496.
PR	XX	14-JUL-2000; 2000US-0218290.
PR	XX	26-JUL-2000; 2000US-0220964.
PR	XX	14-AUG-2000; 2000US-0224518.
PR	XX	14-AUG-2000; 2000US-0224519.
PR	XX	14-AUG-2000; 2000US-0225213.
PR	XX	14-AUG-2000; 2000US-0225214.
PR	XX	14-AUG-2000; 2000US-0225266.
PR	XX	14-AUG-2000; 2000US-0225267.
PR	XX	14-AUG-2000; 2000US-0225268.
PR	XX	14-AUG-2000; 2000US-0225270.
PR	XX	14-AUG-2000; 2000US-0225447.
PR	XX	14-AUG-2000; 2000US-0225757.
PR	XX	14-AUG-2000; 2000US-0225758.
PR	XX	14-AUG-2000; 2000US-0225759.
PR	XX	18-AUG-2000; 2000US-0226279.
PR	XX	22-AUG-2000; 2000US-0226681.
PR	XX	22-AUG-2000; 2000US-0226868.
PR	XX	22-AUG-2000; 2000US-0227182.
PR	XX	23-AUG-2000; 2000US-0227009.
PR	XX	30-AUG-2000; 2000US-0228924.
PR	XX	01-SEP-2000; 2000US-0229287.
PR	XX	01-SEP-2000; 2000US-0229343.
PR	XX	01-SEP-2000; 2000US-0229344.
PR	XX	01-SEP-2000; 2000US-0229345.
PR	XX	05-SEP-2000; 2000US-0229509.
PR	XX	05-SEP-2000; 2000US-0229513.
PR	XX	06-SEP-2000; 2000US-0230437.
PR	XX	06-SEP-2000; 2000US-0230438.
PR	XX	08-SEP-2000; 2000US-0231242.
PR	XX	08-SEP-2000; 2000US-0231243.
PR	XX	08-SEP-2000; 2000US-0231244.
PR	XX	08-SEP-2000; 2000US-0231413.
PR	XX	08-SEP-2000; 2000US-0231414.
PR	XX	08-SEP-2000; 2000US-0232080.
PR	XX	08-SEP-2000; 2000US-0232081.
PR	XX	12-SEP-2000; 2000US-0231968.
PR	XX	14-SEP-2000; 2000US-0232397.
PR	XX	14-SEP-2000; 2000US-0232398.
PR	XX	14-SEP-2000; 2000US-0232399.
PR	XX	14-SEP-2000; 2000US-0232400.
PR	XX	14-SEP-2000; 2000US-0232401.
PR	XX	14-SEP-2000; 2000US-0233063.
PR	XX	14-SEP-2000; 2000US-0233064.
PR	XX	14-SEP-2000; 2000US-0233065.
PR	XX	21-SEP-2000; 2000US-0234223.
PR	XX	21-SEP-2000; 2000US-0234274.
PR	XX	25-SEP-2000; 2000US-0234997.
PR	XX	25-SEP-2000; 2000US-0234998.
PR	XX	26-SEP-2000; 2000US-0235484.
PR	XX	27-SEP-2000; 2000US-0235834.
PR	XX	27-SEP-2000; 2000US-0235835.
PR	XX	29-SEP-2000; 2000US-0236327.
PR	XX	29-SEP-2000; 2000US-0236327.
PR	XX	29-SEP-2000; 2000US-0236367.
PR	XX	05-DEC-2000; 2000US-0256719.
PR	XX	06-DEC-2000; 2000US-0251479.
PR	XX	08-DEC-2000; 2000US-0251856.
PR	XX	08-DEC-2000; 2000US-0251868.
PR	XX	08-DEC-2000; 2000US-0251869.
PR	XX	08-DEC-2000; 2000US-0251989.
PR	XX	08-DEC-2000; 2000US-0251990.
PR	XX	11-DEC-2000; 2000US-0254097.
PR	XX	05-JAN-2001; 2001US-0259678.
XX	XX	(HUMA-) HUMAN GENOME SCI INC.
XX	XX	Rosen CA, Barash SC, Ruben SM;
PI	XX	WPI; 2001-541565/60.
DR	XX	
XX	XX	
XX	XX	
PT	PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT	PT	useful for preventing, diagnosing and/or treating nervous system
PT	PT	cancers and metastases -
XX	XX	
XX	XX	Disclosure; SEQ ID NO 11358; 1701pp + Sequence Listing; English.
XX	XX	
XX	XX	The invention relates to novel genes (ABAl1004-ABAl1534) and proteins
CC	CC	(ABBl14678-ABBl18001) useful for preventing, treating or ameliorating
CC	CC	medical conditions e.g. by protein or gene therapy. The genes are
CC	CC	isolated from a range of human tissues disclosed in the specification.
CC	CC	

PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 metastasis -
 XX
 PS Disclosure; SEQ ID NO 30388; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 5666 BP; 1450 A; 1247 C; 1332 G; 1637 T; 0 other;

 Query Match 3.0%; Score 35.8; DB 22; Length 5666;
 Best Local Similarity 65.8%; Pred. No. 2.1;
 Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

 QY 525 AACGCTGGGAAGAGGCACCTGCCATTACACATCTCCGCTATTCAGGTGTTCTGTC 584
 DB 4528 AATTCCTAGACAGACACCTTCCAGGAACACAGCTCCGCCACTCACTGGTGACCTTC 4469
 QY 585 TGCTGACGACGACGTGCTT 603
 DB 4468 CTCCTGGGAGACTGCTT 4450

 RESULT 39
 AAD06413
 ID AAD06413 standard; DNA; 1614 BP.
 XX
 AC AAD06413;
 XX
 DT 10-AUG-2001 (first entry)
 XX
 DE Human DNAX cytokine receptor subunit 3.2 reverse translational DNA.
 XX
 KW Human; immunomodulator; DNAX cytokine receptor subunit 3.2; DCRS3.2;
 KW therapy; immunological disorder; drug screening; cell development;
 XX chromosome 16p12; ds.
 OS Homo sapiens.
 XX
 PN WO200136467-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 16-NOV-2000; 2000WO-US31363.
 XX
 PR 18-NOV-1999; 99US-0443060.
 PR 13-DEC-1999; 99US-0170320.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Gorman DM;
 XX
 DR WPI; 2001-343800/36.
 XX
 PT New mammalian receptor proteins related to cytokine receptors, useful
 PT for regulating cell development and for diagnosis and treatment of

Db 847 GCCCTCCCTGATGCTTGGAGGACACACACACGCTGCAAGCTG 805

RESULT 43

AAC21988/c
ID AAC21988 standard; cDNA; 355 BP.

AC AAC21988;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 26063.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 1; SEQ ID 26063; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SQ Sequence 355 BP; 90 A; 57 C; 68 G; 119 T; 11 other;

Query Match 2.9%; Score 34.6; DB 21; Length 355;
Best Local Similarity 48.4%; Pred. NO. 1;
Matches 77; Conservative 6; Mismatches 76; Indels 0; Gaps 0;

QY 139 CCGGAACCCGCTGAACAAATTCGGCTGAAGAGGATCCGCCGTATCTGTTCATT 198

DB 329 CTGAATAACAAGAGCTCTCAATTGAGGGGGAAGAGCCCTCCVACTGACATCTARRAT 270

QY 199 TCCCTTTAGCTGACAGACACAATGATCTGTGCGCTTCTGTATATCAAAAC 258

DB 269 TCCTCTTAGACACACAGTACTKCCAAACAACATCTTTCTKGTTTAATGATTAAC 210

QY 259 GGTACTCAATATCTTCTGCGCTGGCTGCATCATCC 297

DB 209 CSRCTGGTCAATCTTTGAAGACTGCCTTCATATAC 171

RESULT 44

AAK81290
ID AAK81290 standard; DNA; 21410 BP.

XX AAK81290;

XX 07-NOV-2001 (first entry)

XX Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:36102.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216847.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0217496.

XX 26-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 18-AUG-2000; 2000US-0225759.

XX 22-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226868.

XX 23-AUG-2000; 2000US-0227009.

XX 30-AUG-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

XX 01-SEP-2000; 2000US-0229344.

XX 01-SEP-2000; 2000US-0229345.

XX 05-SEP-2000; 2000US-0229509.

XX 05-SEP-2000; 2000US-0229513.

XX 06-SEP-2000; 2000US-0230437.

XX 06-SEP-2000; 2000US-0230438.

XX 08-SEP-2000; 2000US-0231242.

XX 08-SEP-2000; 2000US-0231243.

XX 08-SEP-2000; 2000US-0231244.

XX 08-SEP-2000; 2000US-0231413.

XX 08-SEP-2000; 2000US-0231414.

XX 08-SEP-2000; 2000US-0232080.

XX 08-SEP-2000; 2000US-0232081.

XX 12-SEP-2000; 2000US-0231968.

XX 14-SEP-2000; 2000US-0232397.

PR	14-SEP-2000;	2000US-Q233399;
PR	14-SEP-2000;	2000US-Q233399;
PR	14-SEP-2000;	2000US-Q233400;
PR	14-SEP-2000;	2000US-Q233400;
PR	14-SEP-2000;	2000US-Q233401;
PR	14-SEP-2000;	2000US-Q233063;
PR	14-SEP-2000;	2000US-Q233064;
PR	14-SEP-2000;	2000US-Q233065;
PR	21-SEP-2000;	2000US-Q234223;
PR	21-SEP-2000;	2000US-Q234223;
PR	21-SEP-2000;	2000US-Q234274;
PR	25-SEP-2000;	2000US-Q234997;
PR	25-SEP-2000;	2000US-Q234998;
PR	26-SEP-2000;	2000US-Q235484;
PR	27-SEP-2000;	2000US-Q235834;
PR	27-SEP-2000;	2000US-Q235836;
PR	29-SEP-2000;	2000US-Q236327;
PR	-29-SEP-2000;	2000US-Q236327;
PR	29-SEP-2000;	2000US-Q236367;
PR	29-SEP-2000;	2000US-Q236368;
PR	29-SEP-2000;	2000US-Q236369;
PR	29-SEP-2000;	2000US-Q236370;
PR	29-SEP-2000;	2000US-Q236371;
PR	02-OCT-2000;	2000US-Q237038;
PR	02-OCT-2000;	2000US-Q237039;
PR	02-OCT-2000;	2000US-Q237040;
PR	13-OCT-2000;	2000US-Q239935;
PR	13-OCT-2000;	2000US-Q239937;
PR	20-OCT-2000;	2000US-Q240960;
PR	20-OCT-2000;	2000US-Q241221;
PR	20-OCT-2000;	2000US-Q241785;
PR	20-OCT-2000;	2000US-Q241786;
PR	20-OCT-2000;	2000US-Q241787;
PR	20-OCT-2000;	2000US-Q241808;
PR	20-OCT-2000;	2000US-Q241809;
PR	20-OCT-2000;	2000US-Q241826;
PR	01-NOV-2000;	2000US-Q244617;
PR	08-NOV-2000;	2000US-Q246474;
PR	08-NOV-2000;	2000US-Q246475;
PR	08-NOV-2000;	2000US-Q246476;
PR	08-NOV-2000;	2000US-Q246477;
PR	08-NOV-2000;	2000US-Q246478;
PR	08-NOV-2000;	2000US-Q246523;
PR	08-NOV-2000;	2000US-Q246523;
PR	08-NOV-2000;	2000US-Q246524;
PR	08-NOV-2000;	2000US-Q246525;
PR	08-NOV-2000;	2000US-Q246526;
PR	08-NOV-2000;	2000US-Q246527;
PR	08-NOV-2000;	2000US-Q246528;
PR	08-NOV-2000;	2000US-Q246532;
PR	08-NOV-2000;	2000US-Q246532;
PR	08-NOV-2000;	2000US-Q246509;
PR	08-NOV-2000;	2000US-Q246610;
PR	08-NOV-2000;	2000US-Q246611;
PR	08-NOV-2000;	2000US-Q246613;
PR	17-NOV-2000;	2000US-Q249217;
PR	17-NOV-2000;	2000US-Q249208;
PR	17-NOV-2000;	2000US-Q249209;
PR	17-NOV-2000;	2000US-Q249210;
PR	17-NOV-2000;	2000US-Q249211;
PR	17-NOV-2000;	2000US-Q249212;
PR	17-NOV-2000;	2000US-Q249213;
PR	17-NOV-2000;	2000US-Q249214;
PR	17-NOV-2000;	2000US-Q249215;
PR	17-NOV-2000;	2000US-Q249217;
PR	17-NOV-2000;	2000US-Q249218;
PR	17-NOV-2000;	2000US-Q249218;
PR	17-NOV-2000;	2000US-Q249244;
PR	17-NOV-2000;	2000US-Q249245;
PR	17-NOV-2000;	2000US-Q249246;
PR	17-NOV-2000;	2000US-Q249265;
PR	17-NOV-2000;	2000US-Q249297;
PR	17-NOV-2000;	2000US-Q249299;
PR	17-NOV-2000;	2000US-Q249300;
PR	01-DEC-2000;	2000US-Q250160;
PR	01-DEC-2000;	2000US-Q250391;
PR	05-DEC-2000;	2000US-Q251030;

	05-DEC-2000; 2000US-0251988.
PR	06-DEC-2000; 2000US-0256719.
PR	06-DEC-2000; 2000US-0251479.
PR	08-DEC-2000; 2000US-0251856.
PR	08-DEC-2000; 2000US-0251868.
PR	08-DEC-2000; 2000US-0251869.
PR	08-DEC-2000; 2000US-0251989.
PR	08-DEC-2000; 2000US-0251990.
PR	11-DEC-2000; 2000US-0254097.
PR	05-JAN-2001; 2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.
FA	Rosen CA, Barash SC, Ruben SM;
Pf	WPI; 2001-483426/52.
DR	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT	useful for preventing, diagnosing and/or treating cancers and
PT	metastasis -
PS	Disclosure; SEQ ID NO 36102; 307lpp + Sequence Listing; English.
SS	AAMK4951 to AAKG4702 encode the human Immune/Haematopoietic Antigen (I)
CC	amino acid sequences given in AAMK4951 to AAMK4951. (I) have cytosolic
CC	activity, and can be used in gene therapy and vaccine production. (I)
CC	proteins and polynucleotides may be used in the prevention, diagnosis and
CC	treatment of diseases associated with inappropriate (I) expression. For
CC	example, they may be used to treat disorders associated with decreased
CC	expression by rectifying mutations or deletions in a patient's genome
CC	that affect the activity of (I) by expressing inactive proteins or to
CC	suppress the patients own production of (I). Additionally, (I)
CC	polynucleotides may be used to produce the secreted (I), by inserting
CC	the nucleic acids into a host cell and culturing the cell to express the
CC	protein. (I) proteins and polynucleotides may be used to prevent,
CC	diagnose and treat immune/hematopoietic-related diseases, especially
CC	cancers and cancer metastases of haematopoietic-derived cells. AAKM64703
CC	to AAKM64704 represent human Immune/Haematopoietic Antigen Genomic
CC	sequences from the present invention. AAKM64942 to AAKM64950 and AAMK2169
CC	represent sequences used in the exemplification of the present invention.
XX	Sequence 21410 BP; 6408 A; 3742 C; 4370 G; 6890 T; 0 other;
SQ	Query Match 2.9%; Score 34.2; DB 22; Length 21410; Best Local Similarity 52.4%; Pred. No. 16; Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
Oy	56 AGCTGAAGTAGGCCTTTCGTGGCGATTAAATCATTTACCAACTGCCATCATAA 115
Dd	9519 AGCTGAGATCATGTGCATTTGCCTTGACCCTGGCACAGAAGTCAAATCCTCATAA 9578
Oy	116 CAATGTGTAATAATTACTTCCTTACCGGAAAACCCCTGAACAATAATTCGGCTGAAAAAG 175
Dd	9579 AAAAAGAAAAGAAATGTGGTAAATACACATATGGAATACTATTTCAGCTTAAAAAAG 9638
Oy	176 GATCCGCCGTTATCTGTTCATT 198
Dd	9639 GAACACTGTGCTATTGTGACAA 9661
RESULT 45	
AAD15256	
ID	AAD15256 standard; DNA; 74962 bp.
XX	AAD15256;
AC	
DT	
DD	15-NOV-2001 (first entry)
DE	Human phosphatase gene.
Kw	Human; tyrosine phosphatase subfamily; cellular process modulation;
Kw	phosphatase; chromosome 14; single nucleotide polymorphism; SNP;
Kw	variant; ds.

XX PF 12-FEB-2001; 2001WO-US04432.
XX PR 14-FEB-2000; 2000US-0182194.
XX PR 11-OCT-2000; 2000US-0685853.
XX PA (PEKE) PE CORP NY.
XX PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX DR WPI; 2001-529910/58.
XX DR P-PSDB; AAE08552.
XX PT New phosphatase peptide for the screening of modulators useful for
XX PT treating related disorders -
XX PS Claim 4; Fig 3; 109pp; English.
XX CC The present invention provides phosphatase peptides, proteins and
XX CC their polynucleotides. The phosphatase proteins are related to the
XX CC protein tyrosine phosphatase subfamily. The peptides of the invention
XX CC are useful for identifying a modulator of phosphatase peptide expression.
XX CC Modulators of phosphatase peptide are useful for treating diseases or
XX CC conditions mediated by a human phosphatase protein. The phosphatase gene
XX CC of the present invention can be expressed in yeast to identify possible
XX CC interactors and substrates. Artificially synthesised enzymes or derived
XX CC peptides can be used to activate or inhibit cellular processes modulated
XX CC by this phosphatase. The present sequence is a human gene encoding
XX CC phosphatase protein. This gene is located on human chromosome 14.
XX SQ Sequence 74962 BP; 21639 A; 14055 C; 14307 G; 24824 T; 137 other;
Query Match 2.9%; Score 34.2; DB 22; Length 74962;
Best Local Similarity 52.4%; Pred. No. 32;
Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 56 AGCTGAAGTAGGCTGTTCTGTCGGTATTAAATGATGACCGTCCCGCTATTAA 115
DB 5571 AGCTGAGATCATGCCATTGACATCCAGCTGGGCAACAGAGAACTCCATCTCAAA 5630
QY 116 CAATGTGATAAATATCTCCGTTACCGGAAACCGCTGAACAAATTCGGCTGAAAAG 175
DB 5631 AAAAGAGAAAAGAAATGTGTAATACACACATTTGGAATACTATTACGCTTAAAAAG 5690
QY 176 GATCCGCGCTATCTGTTGCATT 198
DB 5691 GAAACTGTGTCATTGTGCAAT 5713
RESULT 46
AAS70590/C
ID AAS70590 standard; cDNA; 645 BP.
XX AC AAS70590;
XX AC
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #5394.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX XN WO200175067-A2.
XX XX 11-OCT-2001.
XX XX 30-MAR-2001; 2001WO-US08631.
XX PF 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX XX

PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX XX WPI; 2001-639362/73.
XX DR P-PSDB; ABG06403.
XX XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 1; SEQ ID No 6394; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 645 BP; 150 A; 182 C; 189 G; 124 T; 0 other;
Query Match 2.9%; Score 34; DB 23; Length 645;
Best Local Similarity 61.1%; Pred. No. 2.3;
Matches 55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 219 AGAGACACAATGATCTGTCGGTCTCTTAATATCAACCGGTACTCAATATCTCTCTG 278
DB 201 AGTGAACACGGCGCTCTCTGTAATGTTGTCACAATAAGTCCAAATATATCTTTCTG 142
QY 279 GCGTGGCTGCCATCATCCGGAAGCGTTCC 308
DB 141 GCACTGGCAGTGGTCACTGCACAGGTCCT 112
RESULT 47
AAC43719/C
ID AAC43719 standard; DNA; 1586 BP.
XX AC AAC43719;
XX AC
XX DT 18-OCT-2000 (first entry)
XX XX Zea mays DNA fragment SEQ ID NO: 40258.
XX DE Zea mays subsp. mays.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic;
XX KW pathway; promoter; termination sequence; corn; ss.
XX OS Zea mays subsp. mays.
XX XN EP1033405-A2.
XX XX 06-SEP-2000.
XX XX 25-FEB-2000; 2000EP-0301439.
XX PF

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

DR 29-SEP-2000; 2000US-0236368.
XX 29-SEP-2000; 2000US-0236369.
PT 29-SEP-2000; 2000US-0236370.
PT 02-OCT-2000; 2000US-0236802.
XX 02-OCT-2000; 2000US-0237037.
XX 02-OCT-2000; 2000US-0237038.
XX 02-OCT-2000; 2000US-0237039.
XX 02-OCT-2000; 2000US-0237040.
XX 13-OCT-2000; 2000US-0239935.
XX 13-OCT-2000; 2000US-0239937.
XX 20-OCT-2000; 2000US-0240960.
XX 20-OCT-2000; 2000US-0241221.
XX 20-OCT-2000; 2000US-0241785.
XX 20-OCT-2000; 2000US-0241786.
XX 20-OCT-2000; 2000US-0241787.
XX 20-OCT-2000; 2000US-0241808.
XX 20-OCT-2000; 2000US-0241809.
XX 20-OCT-2000; 2000US-0241826.
XX 01-NOV-2000; 2000US-0244617.
XX 08-NOV-2000; 2000US-0246474.
XX 08-NOV-2000; 2000US-0246475.
XX 08-NOV-2000; 2000US-0246476.
XX 08-NOV-2000; 2000US-0246477.
XX 08-NOV-2000; 2000US-0246478.
XX 08-NOV-2000; 2000US-0246523.
XX 08-NOV-2000; 2000US-0246524.
XX 08-NOV-2000; 2000US-0246525.
XX 08-NOV-2000; 2000US-0246526.
XX 08-NOV-2000; 2000US-0246527.
XX 08-NOV-2000; 2000US-0246528.
XX 08-NOV-2000; 2000US-0246532.
XX 08-NOV-2000; 2000US-0246609.
XX 08-NOV-2000; 2000US-0246610.
XX 08-NOV-2000; 2000US-0246611.
XX 08-NOV-2000; 2000US-0246613.
XX 17-NOV-2000; 2000US-0249207.
XX 17-NOV-2000; 2000US-0249208.
XX 17-NOV-2000; 2000US-0249209.
XX 17-NOV-2000; 2000US-0249210.
XX 17-NOV-2000; 2000US-0249211.
XX 17-NOV-2000; 2000US-0249212.
XX 17-NOV-2000; 2000US-0249213.
XX 17-NOV-2000; 2000US-0249214.
XX 17-NOV-2000; 2000US-0249215.
XX 17-NOV-2000; 2000US-0249216.
XX 17-NOV-2000; 2000US-0249217.
XX 17-NOV-2000; 2000US-0249218.
XX 17-NOV-2000; 2000US-0249244.
XX 17-NOV-2000; 2000US-0249245.
XX 17-NOV-2000; 2000US-0249264.
XX 17-NOV-2000; 2000US-0249265.
XX 17-NOV-2000; 2000US-0249297.
XX 17-NOV-2000; 2000US-0249299.
XX 17-NOV-2000; 2000US-0249300.
XX 01-DEC-2000; 2000US-0250160.
XX 01-DEC-2000; 2000US-0250391.
XX 05-DEC-2000; 2000US-0251030.
XX 05-DEC-2000; 2000US-0251988.
XX 05-DEC-2000; 2000US-0256719.
XX 08-DEC-2000; 2000US-0251479.
XX 08-DEC-2000; 2000US-0251856.
XX 08-DEC-2000; 2000US-0251868.
XX 08-DEC-2000; 2000US-0251869.
XX 08-DEC-2000; 2000US-0251989.
XX 08-DEC-2000; 2000US-0251990.
XX 11-DEC-2000; 2000US-0254097.
XX 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX

P-PSDB; AAM87436.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
Claim 1; SEQ ID NO 5277; 307lpp + Sequence Listing; English.
AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.
XX Sequence 333 BP; 71 A; 88 C; 77 G; 93 T; 4 other;
SQ

Query Match 2.9%; Score 33.8; DB 22; Length 333;
Best Local Similarity 51.8%; Pred. No. 1.9;
Matches 71; Conservative 2; Mismatches 64; Indels 0; Gaps 0;
QY 141 GGAAGAGATAGAGGGAAGGTGGTGGCAATCAGACTGGACAGCTCATGTCCCTCAGC 158
DB 99 GGAAGAGATAGAGGGAAGGTGGTGGCAATCAGACTGGACAGCTCATGTCCCTCAGC 158
QY 201 CCTTTAGCCCTGACTAGCCAGACACAATCATCTGTGCCGTTCTGTTAATATCAAAACCGG 260
DB 159 CACTTGGGCTGRTGGGAAAGAACACATCTTCAGTATGGTTCTTCTCTACTAG 218
QY 261 TACTCAATATCTTCTCT 277
DB 219 CCTCTTCATCTCATCT 235

RESULT 49
ABV48867
ID ABV48867 standard; cDNA; 474 BP.
XX AC ABV48867;
XX 17-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 48858.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX Pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -

XX Claim 1; Page 9566-9567; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 474 BP; 140 A; 114 C; 118 G; 102 T; 0 other;

Query Match 2.9%; Score 33.8; DB 23; Length 474;

Best Local Similarity 55.68; Pred. No. 2.3;

Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 7 GAGATGGAAAAAGCCAAATAAATAATGCCCCATCCAGCGCGCTCCAGCTGAAGA 66

DB 238 GAAATAAAAAAAGAAAAAAGGCGCGCCGCCAGCGGTGGAGCTCCAATTC 297

QY 67 GGCGTGTCTGCGGTATTAAATGCAATGACCGTCCCGGTATTAAACAATGTGA 123

DB 298 GCCTATAGTAGTCGTATTACGCGCGCTCACTGGCGGTGTTTTACAGGTCGTGA 354

RESULT 50

AAQ21621/C

ID AAQ21621 standard; DNA; 1647 BP.

XX AAQ21621;

XX 29-JUN-1992 (first entry)

XX Human Tryptophan-2,3-dioxygenase coding sequence HTO3.

XX Tourette Syndrome; TD02; kyneurenine formation; autism; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 34..1344

XX FT /*tag= a

XX CDS 34..1008

XX FT /*tag= b

XX WO9202637-A.

XX 20-FEB-1992.

XX 02-AUG-1991; 91WO-US05441.

XX 14-JUN-1991; 91US-0715660.

XX 03-AUG-1990; 90US-0562596.

XX (CITY) CITY OF HOPE.

XX Comings DE;

XX WPI; 1992-080090/10.

XX P-PSDB; AAR21547.

XX Diagnosis of Tourette syndrome and associated disorders - by
PT detecting mutation in tryptophan oxygenase-indole:amine oxygenase
PT genes and prods.

XX Disclosure; Fig 5A; 45pp; English.

XX Clone HTO3 was isolated from a human liver mRNA-derived cDNA
CC library screened with a rat liver tryptophan oxygenase cDNA probe
CC (see AAQ21622).

CC The specification includes two different sequences which are

CC both described as "HTO3" (see also AAQ21620). They are identical

CC through the first 741 nucleotides of AAQ21620 before the much longer

CC sequence AAQ21621 begins to diverge. The human TD02 gene sequence,

CC including regulatory sequences, is also given in the specification,

CC as is a genomic TD02 sequence which includes the introns (see AAQ21623

CC and AAQ21624, respectively). The sequence of the open reading frame in

CC these sequences differs from AAQ21621 (and AAQ21620) by the deletion of

CC a single adenine residue (i.e. from the run of 8 A's at position

CC 713-720 of AAQ21621). Yet another Figure, which illustrates the

CC degree of homology between HTO3 and the rat TD02 sequence, also

CC lacks the 8th Adenine, suggesting that the "extra" A in both

CC AAQ21620 and AAQ21621 represents a typographical error. The patent

CC specification does not contain any discussion of the apparent

CC differences and includes only the minimum description of each of

CC the sequences. Consequently, the significance of an in frame

CC termination codon within the longest CDS of HTO3 is not known.

XX Sequence 1647 BP; 556 A; 292 C; 327 G; 472 T; 0 other;

Query Match 2.9%;

Best Local Similarity 53.48; Pred. No. 4.7; Length 1647;

Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1013 ATTACTGTCTGCGCACTATGGCAGATGACCCAGATCAGGTTTAAATCCCGATAATCCG 1072

DB 345 ATCTCTCACATGGCCATTCTGAAAGATCTCTCGAACACAGAAATCCCACTCCCGAGAGATTG 286

QY 1073 TCGAAGTCTGAGGATGGAAGGAAGCTGTAAGCTTCTTCTGAAAGATATAAAGTGCACATC 1132

DB 285 CTTAAACCCAGAGTTTCATAAGCTTGATGAGTTATGATAAAAGATGTTTCATCATGGATTTT 226

QY 1133 ATGCCCTCTTTT 1145

DB 225 ATTTCCCTTTGTT 213

Search completed: June 18, 2003, 02:09:39

Job time : 223.603 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 01:20:59 ; Search time 1249.12 Seconds
(without alignments)
15312.305 Million cell updates/sec

Title: US-09-674-277-2

Perfect score: 1181

Sequence: 1 ctcgaggaatggaagaaaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48.8	4.1	755	17	AQ989708
2	46.6	3.9	615	10	AV909088
3	46.6	3.9	629	13	BJ264853
4	46.6	3.9	669	10	AV910971
5	46.6	3.9	693	10	AV945422
6	45.8	3.9	647	13	BJ312175

495	14	BQ763870	BQ763870	EBR003_SQ
714	9	AU100720	AU100720	
866	17	BH468052	BH468052	BOHOJ24TF
1101	17	CNS0106X	CNS0106X	AL098595 Drosophila
533	11	AW042265	AW042265	614026D01
736	11	AY112421	AY112421	zea mays
645	14	BQ087067	BQ087067	Cri_9_106
461	17	BH778732	BH778732	fzmb013f0
769	17	BH388561	BH388561	AG-ND-130
926	12	BG330039	BG330039	602428912
509	13	BI744669	BI744669	xk89h09.y
632	12	BG605346	BG605346	WHE2331.F
479	13	BJ292385	BJ292385	RJ292385
545	14	C20457	C20457	Rice
756	17	CNS02D50	CNS02D50	Tetraodon
649	10	AV784670	AV784670	AV784670
501	12	BG840188	BG840188	MEST8-G12
540	12	BF586903	BF586903	FM1_31_H0
672	13	BI968888	BI968888	GM830006B
499	13	BI498039	BI498039	sag15c08.
516	13	BM143281	BM143281	sa142b12.
609	10	AV836868	AV836868	AV836868
849	17	CNS031JD	CNS031JD	Tetraodon
974	14	BQ878757	BQ878757	AGENCOURT
948	17	CNS017IM	CNS017IM	Drosophila
320	10	AW453731	AW453731	zeh10506
831	12	BG565743	BG565743	602589078
582	10	AV987761	AV987761	AV987761
1101	17	CNS00LT2	CNS00LT2	AL078714 Drosophila
861	12	BG338455	BG338455	602587161
432	10	AW509231	AW509231	sh92h11.y
445	14	BQ611687	BQ611687	sap63g01.
446	10	AW459606	AW459606	sh89c12.y
478	9	AI496532	AI496532	sh13a03.y
507	10	AW620859	AW620859	sj47a11.y
512	17	AQ785148	AQ785148	HS_3159.A
514	17	AQ662456	AQ662456	HS_5455.A
523	10	BE440931	BE440931	sp49g12.y
527	10	AW279073	AW279073	sg07d10.y
540	13	BM092322	BM092322	sah11e10.y
542	13	BI427174	BI427174	sah77c01.
566	13	BI427022	BI427022	sah77c01.
581	14	BQ530772	BQ530772	sap31a10.
584	13	BI787127	BI787127	sag71f10.
585	10	AW397142	AW397142	sg67d08.y
586	14	BQ629541	BQ629541	saq03d02.
636	10	AW164393	AW164393	se71h10.y
821	12	BG618482	BG618482	602645674
1101	17	CNS014TP	CNS014TP	AL104599 Drosophila
319	9	AI229512	AI229512	EST226207
482	12	BF565804	BF565804	UT-R-C4-a
557	14	BM860712	BM860712	fy53f12.x
248	10	AV358623	AV358623	AV358623
254	17	AQ078429	AQ078429	CIT-HSP-2
372	12	BF404709	BF404709	UT-R-CAL-
412	10	BB699898	BB699898	BB699898
434	13	BB699994	BB699994	BB699994
438	13	BJ449621	BJ449621	BJ449621
520	10	AV906979	AV906979	AV906979
523	13	BJ450987	BJ450987	BJ450987
545	14	BP017181	BP017181	BP017181
600	13	BJ458496	BJ458496	BJ458496
650	10	BE279974	BE279974	601156988
406	17	AQ184201	AQ184201	HS_3221.A
646	14	BQ562752	BQ562752	HA077G06-
663	10	BB662428	BB662428	BB662428
972	11	BQ715820	BQ715820	AGENCOURT
3321	14	AK014488	AK014488	Mus muscu
576	17	AQ231338	AQ231338	HS_2053.A
616	9	AL635769	AL635769	AL635769
712	2	BF273334	BF273334	GA_EB001
793	13	BI821211	BI821211	603034921
925	17	CNS0091P	CNS0091P	AL053013 Drosophila

C 80 34 1081 12 BF341523
 C 81 33.8 2.9 423 17 AZ373715
 C 82 33.8 2.9 533 17 BG565053
 C 83 33.8 2.9 577 13 BE020591
 C 84 33.8 2.9 539 13 RI833613
 C 85 33.8 2.9 583 10 AV647553
 C 86 33.8 2.9 626 10 AV655093
 C 87 33.8 2.9 627 12 BG562641
 C 88 33.8 2.9 630 10 AV655934
 C 89 33.8 2.9 630 12 BG545282
 C 90 33.8 2.9 651 12 BG568315
 C 91 33.8 2.9 653 12 BG565518
 C 92 33.8 2.9 660 12 BG565394
 C 93 33.8 2.9 665 12 BG568577
 C 94 33.8 2.9 666 12 BE728670
 C 95 33.8 2.9 672 10 AV684416
 C 96 33.8 2.9 719 12 BG568953
 C 97 33.8 2.9 719 17 AZ463561
 C 98 33.8 2.9 732 12 BG564876
 C 99 33.8 2.9 740 10 AV653890
 C 100 33.8 2.9 747 10 AV648964
 C 101 33.8 2.9 756 12 BG563207
 C 102 33.8 2.9 759 13 RI834454
 C 103 33.8 2.9 761 12 BG618475
 C 104 33.8 2.9 763 12 BG567109
 C 105 33.8 2.9 780 12 BG563303
 C 106 33.8 2.9 785 12 BG564030
 C 107 33.8 2.9 792 12 BG616785
 C 108 33.8 2.9 797 12 BG562824
 C 109 33.8 2.9 809 12 BG564329
 C 110 33.8 2.9 812 12 BG617332
 C 111 33.8 2.9 818 9 AU120839
 C 112 33.8 2.9 836 12 BF126698
 C 113 33.8 2.9 851 12 BG568106
 C 114 33.8 2.9 869 12 BG533516
 C 115 33.8 2.9 891 10 BE380004
 C 116 33.8 2.9 955 12 BG617602
 C 117 33.8 2.9 1002 17 CNS017AP
 C 118 33.8 2.9 1032 17 CNS07B6X
 C 119 33.8 2.9 1085 17 CNS0409L
 C 120 33.6 2.8 701 17 AQ110329
 C 121 33.6 2.8 713 9 AU056834
 C 122 33.6 2.8 736 12 BF860318
 C 123 33.6 2.8 797 17 CNS003F8
 C 124 33.6 2.8 939 17 CNS00CNG
 C 125 33.6 2.8 1065 12 BG026002
 C 126 33.4 2.8 323 9 AA196883
 C 127 33.4 2.8 372 9 AI858306
 C 128 33.4 2.8 466 13 BI961758
 C 129 33.4 2.8 490 9 AI743392
 C 130 33.4 2.8 654 12 BG505886
 C 131 33.4 2.8 800 12 BF248131
 C 132 33.4 2.8 814 12 BG495140
 C 133 33.4 2.8 972 14 BQ948291
 C 134 33.2 2.8 427 17 AZ049336
 C 135 33.2 2.8 453 13 BM430341
 C 136 33.2 2.8 490 17 BH871177
 C 137 33.2 2.8 553 10 AW658701
 C 138 33.2 2.8 636 13 BM359682
 C 139 33.2 2.8 642 12 BG443931
 C 140 33.2 2.8 719 17 BH360392
 C 141 33.2 2.8 730 17 AQ746896
 C 142 33.2 2.8 777 12 BF244695
 C 143 33.2 2.8 932 17 CNS0298Y
 C 144 33.2 2.8 980 12 BG106295
 C 145 33.2 2.8 980 17 CNS041NY
 C 146 33.2 2.8 1024 17 CNS00RHJ
 C 147 33 2.8 293 12 BG892595
 C 148 33 2.8 317 10 AW530608
 C 149 33 2.8 334 10 AW530607
 C 150 33 2.8 388 13 BJ458049

ALIGNMENTS

RESULT 1
LOCUS

DEFINITION

AQ989708 755 bp DNA linear GSS 14-AUG-2000
 Rf000285 Photorhabdus luminescens strain W14 M13 library
 Photorhabdus luminescens genomic clone PLG00285, DNA sequence.

ACCESSION

AQ989708
 AQ989708.1 GI:9648302

KEYWORDS

SOURCE

Photorhabdus luminescens.
 Photorhabdus luminescens
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Photorhabdus.

REFERENCE

1 (bases 1 to 755)
 french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
 Daborn,P.J., Bowen,D. and Blattner,F.R.

AUTHORS

A genomic sample sequence of the entomopathogenic bacterium
 Photorhabdus luminescens W14: potential implications for virulence

TITLE

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: french-Constant RH
 Department of Biology and Biochemistry
 University of Bath
 South Building, Bath BA2 7AY, UK
 Tel: (44) 1225 826621
 Fax: (44) 1225 826779
 Email: bsr@bath.ac.uk

This is one of 2,122 random reads from the M13 library. For
 annotation of identified clones (BLASTX, BLASTN and mapping to E.
 coli K12 genome) please see french-Constant et al. 2000, Nucleic
 Acids Res.

Seq primer: M13 Forward
 Class: shotgun.

Location/Qualifiers
 i. .755

FEATURES

source

/organism="Photorhabdus luminescens"
 /strain="W14"
 /db_xref="taxon:29488"
 /clone="PLG00285"
 /clone_lib="Photorhabdus luminescens strain W14 M13
 library"

/dev_stage="primary phase variant"
 /note="Genomic DNA from strain W14 was size selected (1-2
 kb) and then cloned into M13 Janus."

BASE COUNT 204 a 194 c 136 g 209 t 12 others

ORIGIN

Query Match 4.1%; Score 48.8; DB 17; Length 755;
 Best Local Similarity 54.5%; Pred. No. 0.011;

Matches 120; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

QY 264 TCATATCTTCTCTGGCTGCTGCCATATCCGGAAGCTTCGGTGGGATAAAAA 323
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 53 TCATATCTTCTCTGGCTGCTGCCATATCCGGAAGCTTCGGTGGGATAAAAA 112
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 324 TCGCGAGTCCGCGCTCCATGCACACATCCCGGTAACAGCTCCCTGTACA 383
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 113 AAATTAACGGCATGACAAATTCCTGGAACATCTCTGGCGTAACCAATACCTGAATTA 172
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 384 TTCTTCTGAATGACA---TCAGGGATCCGCGCTCACTGGCGATACGGGACGCG 440
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 173 CTTTCATACCGACAACTCCGGTATCCGATATACCTGGCAATTAACCGCTGCCA 232
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 441 GAGACTGAGCTTACGCCAGTACCATACCAACGCTTCAT 480
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 233 CAGCCATGCTTCGGCAATCGTAATCCCAACGCTTCGT 272
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
 AV909088/c AV909088 615 bp mRNA linear EST 18-JAN-2002

LOCUS

DEFINITION	AV090888 K. Sato unpublished cDNA library, cv. Akashinriki vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone baak11122 5', mRNA sequence.	
ACCESSION	AV090888	
VERSION	AV090888.1 GI:18204518	
KEYWORDS	EST.	
SOURCE	Hordeum vulgare subsp. vulgare.	
ORGANISM	Hordeum vulgare subsp. vulgare	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.	
AUTHORS	Sato,K., Saitoh,D. and Takeda,K.	
TITLE	Barley EST sequencing project in NIG and Okayama Univ	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6836 Fax: 81-559-81-6855 Email: tshinigenes.nig.ac.jp.	
FEATURES	Location/Qualifiers	
SOURCE	1..615	
	/organism="Hordeum vulgare subsp. vulgare"	
	/cultivar="Akashinriki"	
	/db_xref="taxon:112509"	
	/clone="baak11122"	
	/clone_lib="K. Sato unpublished cDNA library, cv. Akashinriki vegetative stage leaves"	
	/tissue_type="leaves"	
	/dev_stage="vegetative stage"	
BASE COUNT	163 a	118 c 193 g 141 t
ORIGIN		
Query Match	3.9%;	Score 46.6; DB 10; Length 615;
Best Local Similarity	56.1%;	Pred. No. 0.043;
Matches	88; Conservative	0; Mismatches 69; Indels 0; Gaps 0;
Qy	371	CGTCCCTGTGCACATTTCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
Db	444	CTTCTCTCTCGTCTCAGGATATATGTCAGTATGCTTCAGCTCGAGCTCGACGACGCCAC 385
Qy	431	GGCAGCGCGGAGCTGACGCTTCAGCCAGTACCATACCAACAGCTTCATTTTCGAAGG 490
Db	384	CGGAAGTCCCGATGACATTCGCTCCCAACCAACCAACCCAGCGTCTCTGATTCAAGAGG 325
Qy	491	CATGACCACCACTGGCAATCCGGTAGACCGGTAAC 527
Db	324	CATCAACAACACATCCCACTGGCATAGGCGCTGTGAC 288
RESULT 3		
BJ264853		
LOCUS	BJ264853 Y. Ogihara unpublished cDNA library, wh_h Triticum aestivum cDNA clone whh9f24 3', mRNA sequence.	
DEFINITION	629 bp mRNA linear EST 08-APR-2002	
ACCESSION	BJ264853	
VERSION	BJ264853.1 GI:20085028	
KEYWORDS	EST.	
SOURCE	bread wheat.	
ORGANISM	Triticum aestivum	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.	
AUTHORS	Ogihara,Y. and Murai,K.	
TITLE	Expressed genes in Triticum aestivum	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan	

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

FEATURES
source Location/Qualifiers

1..866
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHOJ24"
/clone_lib="BOHO"
/note="vector: pHS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS1 using BstXI linkers"

BASE COUNT 283 a 200 c 144 g 239 t

ORIGIN
Query Match 3.5%; Score 41.6; DB 17; Length 866;
Best Local Similarity 56.6%; Pred. No. 1.2;
Matches 77; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 391 GAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGCACCGCGAGACTGACG 450
DB 116 GGATGATATCAGGATTCACCGCCGAGCCGCAACACAGGAGTCCAGAGCCATTG 175
QY 451 CTTACGACGATACCATACCAACGCTTCATTTTCGAGGATGACACACACACTGGCAA 510
DB 176 CTTCAAGACACACAGGCGAGCTCTCGACTCTGATGGATCATCAACACATCTCCG 235
QY 511 TCCGCTAGACCGGTAA 526
DB 236 TTGCGTAAGCTGTGA 251

RESULT 10
CNS0106X 1101 bp DNA linear GSS 26-JUL-1999
LOCUS
DEFINITION
BACH03K20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION
AL098595 1 GI:5610206
VERSION
GSS.
KEYWORDS
Drosophila melanogaster.
SOURCE
Drosophila melanogaster.
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACH03K20"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

BASE COUNT 258 a 107 c 60 g 175 t 501 others
ORIGIN
Query Match 3.5%; Score 41; DB 17; Length 1101;

Best Local Similarity 16.7%; Pred. No. 1.8;
Matches 42; Conservative 118; Mismatches 91; Indels 0; Gaps 0;
QY 881 GGGTATAGGACACAAATCAATCAGGGAACGTTCTGGTCTT 940
DB 742 KDKAKAKADARDKRRWDKDKKRAKAAKADADADADAKADAKAKADADA 801
QY 941 ACGGGTATAGTGGTTTGTCTGACATAGTGAAGCGGTGACACGATATCAGACGCT 1000
DB 802 DGGGDKKKRAKDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 861
QY 1001 CAGTCTCTGATATATTACTGTCTATGCGCCATGAGGAGATGACGATAGTTTAAATC 1060
DB 862 AKADAKAAKAKAKADDDAKAATKAKATKAKAKAKAKKKKKKKKKKKKKKKKK 921
QY 1061 CCGGATAATCCGTCGAGTCTGAGGATGGAAGGCTGAGGCTGTTCTCTCAAGGAAT 1120
DB 922 KDDKDDKDDKDDKADKDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 981
QY 1121 AAAAGTGACAT 1131
DB 982 DADAKAKAKAK 992

RESULT 11
AW042265 533 bp mRNA linear EST 15-SEP-1999
LOCUS
DEFINITION
614026D01.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA, mRNA sequence.
ACCESSION
AW042265
VERSION
AW042265.1 GI:5901165
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 533)
Walbot, V.
AUTHORS
Maize ESTs from various cDNA libraries sequenced at Stanford University
TITLE
Unpublished (1999)
JOURNAL
Contact: Walbot V
COMMENT
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614026 row: D column: 01.
Location/Qualifiers
1..533
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XfOLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site: 1; EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)";

BASE COUNT 148 a 108 c 163 g 114 t
ORIGIN
Query Match 3.4%; Score 40.6; DB 10; Length 533;
Best Local Similarity 55.2%; Pred. No. 2.1;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 384 TTCTTCTCAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGCACCGCGAG 443
DB 301 TCCTTGGGTATATATATATAGGTATCCCTCCACAGCAGCGACGACTGGGACTCCAGAA 242


```

QY 444 ACTGAGCTTCAGCCAGTACCATACCAACAGCTTCATTTCCTCCAGAGGCATGACCACCACA 503
Db 241 GCATGGACTCAGCACTACTTGCCTGAGGGTCTCGGATTCGAAGGCATTCGAATACA 182
QY 504 CTGGCAATCCGGTAGACCGGTAA 526
Db 181 TCCGCACTGCGTACGCTTGTA 159

RESULT 12
AY112421
LOCUS Zea mays CL8980_1 mRNA sequence. 736 bp mRNA linear HTC 26-MAY-2002
DEFINITION
ACCESSION AY112421
VERSION AY112421.1 GI:21217011
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 736)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 736)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
source
1..736
/organism="Zea mays"
/db_xref="MaizeDB:633031"
/db_xref="taxon:4577"
/clone="CL8980_1"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project."
BASE COUNT 171 a 198 c 154 g 182 t 31 others
ORIGIN
Query Match 3.4%; Score 40.6; DB 11; Length 736;
Best Local Similarity 55.2%; Pred. No. 2.2;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 384 TTCTTCTGATGACATCAGGATCCGCGCTCTACTGGCGATACAGCGGACGCGGAG 443
Db 438 TCCTTGGGTATATATACAGTATCCCTCCAGCAGCAGCAGCAGCTGGACTCCAGAA 497
QY 444 ACTGAGCTTCAGCCAGTACCATACCAACAGCTTCATTTCCTCCAGAGGCATGACCACCACA 503
Db 498 GCCATGGACTCCAGCACTACTTGCCTGAGGGTCTCGGATTCGAAGGCATTCGAATACA 557
QY 504 CTGGCAATCCGGTAGACCGGTAA 526
Db 558 TCCGCACTGCGTACGCTTGTA 580

RESULT 13
BQ087067/c
LOCUS Zea mays CL8980_1 mRNA sequence. 645 bp mRNA linear EST 05-APR-2002
DEFINITION
ACCESSION BQ087067
VERSION BQ087067.1 GI:20046268

```

```

KEYWORDS Ceratopteris richardii.
SOURCE Ceratopteris richardii
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Filicophyta; Filicopsida; Filicales; Pteridaceae; Ceratopteris.
1 (bases 1 to 645)
Chatterjee,A., San Miguel,P., Stout,S.C., Banks,J. and Roux,S.J.
Expressed sequence tags of cDNA clones from a C. richardii library
Unpublished (2000)
CONTACT: Roux SJ
Section of Molecular Cell and Developmental Biology
University of Texas
Biology Building, Room 16, Austin, TX 78712, USA
Tel: 512 471 4238
Fax: 512 232 3402
Email: sroux@uts.cc.utexas.edu
Plate: Cri_9 row: I column: 06
Seq primer: SP6.
FEATURES
source
1..645
/organism="Ceratopteris richardii"
/cultivar="Brogan"
/db_xref="taxon:49495"
/clone="Cri_9_I06"
/clone_lib="Ceratopteris Spore Library"
/tissue_type="Gametophyte"
/cell_type="Spore"
/dev_stage="20 hours after germination initiation"
/notes="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."
BASE COUNT 159 a 128 c 186 g 171 t 1 others
ORIGIN
Query Match 3.3%; Score 39.4; DB 14; Length 645;
Best Local Similarity 53.8%; Pred. No. 4.7;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 374 CCTGTGCATCTTCTGATGACATCAGGATCCGCGCTCTACTGGCGATACAGGG 433
Db 535 CCCTCTCTGATTCGCTGGGATATATCAGGAATGCCCTGCTTCTGCTCCCAATGG 476
QY 434 CAGCGCGGAGACTGACCTTCACCCAGTACCATACCAAGCGCTTCATTTCCGAAGCAT 493
Db 475 TACCCCTGATGCCATAGCCTCAAGAACACAGCAAGAGTTTCTGATTTCAGAGGGTGT 416
QY 494 GACCAACACATCGGTAATCCGCTAGACCGGTAA 526
Db 415 CACAAAACATCCCACTTGCATAGGCTTGTA 383

RESULT 14
BH778732/c
LOCUS Zea mays 461 bp DNA linear GSS 28-MAR-2002
DEFINITION fzm013f051a07f0 fzm6 filtered library Zea mays genomic clone
fzm013f051a07 5', DNA sequence.
ACCESSION BH778732
VERSION BH778732.1 GI:19781391
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 461)
Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
Generalthresh methylation filtered genomic sequences from maize
Unpublished (2002)
CONTACT: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979

```

Thu Jun 19 10:21:33 2003

Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: f2mb013f051 row: a column: 07
Seq primer: M13 forward
Class: shotgun
High quality sequence stop: 461.

FEATURES source

Location/Qualifiers
1. 461
/organism="Zea mays"
/cultivar="M017"
/db_xref="taxon:4577"
/clone="f2mb013f051a07"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HindIII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HindIII-digested pBCSK(-) vector and electroporated into E. coli cells."
122 a 98 c 120 g 121 t

BASE COUNT 122 a 98 c 120 g 121 t

Query Match 3.3%; Score 39.2; DB 17; Length 461;
Best Local Similarity 53.2%; Pred. No. 5.1;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 371 CGTCCCTGCATCTTCTGAATGACATCAGGATCCCGCGTCTCACTGGCGATAAC 430
DB 399 CTTACCTCTCTGTGCTTCGGTATATATCAGGTATCCCTCCAGCAGCAGCAGC 340
QY 431 GGGCAGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTATTTCCGAAGG 490
DB 339 TGGGATCCAGACCATGGACTCCAGCACTATTTGCCGAGGTCTCGGATTCTGAAGG 280
QY 491 CATGACACACATCGGATCCGGTAGACCGGTAA 526
DB 279 CATTGCAATACATCTCCACTGGCGTACGCTTGTA 244

RESULT 15
BH388561

LOCUS AG-ND-130K10.TF ND-TAM Anopheles gambiae genomic clone AG-ND-130K10
DEFINITION DNA sequence.

ACCESSION BH388561
VERSION BH388561.1 GI:17334702

KEYWORDS GSS
SOURCE African malaria mosquito.

ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.

REFERENCE 1 (bases 1 to 769)
Shetty J., Malek J., Koo H., Collins F., Gardner M. and Loftus B.J.
Direct Submission of BAC-end sequences from Anopheles gambiae
Unpublished (2001)
Other_GSSs: AG-ND-130K10.TF

CONTACT: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208
Fax: 301 838 3543

Email: bjoftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain (TIGR). All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For

Class: BAC ends.

FEATURES source

Location/Qualifiers
1. 769
/organism="Anopheles gambiae"
/strain="PEST"
/db_xref="taxon:7165"
/clone="AG-ND-130K10"
/clone_lib="ND-TAM"
/note="Vector: pECBAC1; Site_1: HindIII"
212 a 157 c 143 g 257 t

BASE COUNT
ORIGIN

Query Match 3.3%; Score 38.6; DB 17; Length 769;
Best Local Similarity 52.1%; Pred. No. 8;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 345 GCAGACATCCCCACGGTAACAGCTCCCTGTCACATTTCTGTGAATGACATCAGG 404
DB 137 GCATCATACGCCAACATCTGCAATAATCCGGTTACTCCGTGTTTGTACTCCGGT 196
QY 405 ATCCGCCCGCTCTCACTGGCGATAACGGCGAGACTGAGCTTACGCCAGTACC 464
DB 197 ATTCGCCCGCATTAATCTACTGATTACCGGAGTTCCGCGAGCATTTCTCCAGAGTCA 256
QY 465 ATACCAAGCGTTCTATTTCCGAAGCATGACCACACACTGGCA 509
DB 257 AGACCGAAGCTTCTCTGTTGTGACGGAAGCATATAATACATCCGCA 301

RESULT 16
BG330039/c

LOCUS 602428912F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:455814 5',
DEFINITION mRNA sequence.

ACCESSION BG330039

VERSION BG330039.1 GI:13136477

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 926)
NIH-MGC <http://mgc.nci.nih.gov/>

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgaps@email.nih.gov

Tissue procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LICM1264 row: h column: 03

High quality sequence stop: 667.

FEATURES

source

Location/Qualifiers
1. 926

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:455814"

/clone_lib="NIH_MGC_16"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:

ECORI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

129 a 287 c 264 g 246 t

BASE COUNT

ORIGIN

	QY	324	TCGCGCAGTGC	CGCGGTCC	ATGCAGACACAT	CCCCACGGGTAA	CAGCGTCCCTG	TACACA	383
Best Local Similarity	52.5%								
Mismatches	83								
Conservative	0								
Indels	75								
Gaps	0								
QY		373	TCAACCAT	TGCGTTCGGC	CAGCTTAT	TCCATCGGTAT	CAGAAATCT	TTCGTCCTG	314
Db									
QY		384	TTCTTCTG	ATGACATC	AGGGATCC	CGCCCGTCT	CATGGCGATAC	CGGCACGCGG	443
Db		313	CTTCAATA	ACTGCTCT	GTGAGGCC	CCACCAC	TCTCAATAC	ATACTG	254
QY		444	ACTGAGCG	TTACGCCAG	TACCAT	TACCAAA	CGCTTCATT	481	
Db		253	TCTAAGCG	TTCAAT	TGGGACA	TATACCAAA	TGTTTCATT	216	

RESULT 18	
BG6053346/c	
LOCUS	632 bp mRNA linear EST 16-APR-2001
DEFINITION	WHE2331_Fll_L20S Wheat pre-anthesis spike cDNA library Triticum aestivum cDNA clone WHE2331_Fll_L20, mRNA sequence.
ACCESSION	BG6053346
VERSION	BG6053346
KEYWORDS	EST.
SOURCE	Bread wheat. Triticum aestivum
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum.

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Croser, J.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat genomes - Pre-anthesis spike cDNA library

JOURNAL COMMENT Unpublished (2000)

Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510/5595773

West Arisa, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510595773
Fax: 510595818
Email: caanders@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: StrataGene SK primer.

clone_lib="Wheat pre-anthesis spike cDNA library"	
tissue_type="Spike before anthesis"	
dev_stage="Adult plant"	
lab_host="E. coli SOLR"	
note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; plants were grown in the greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."	
145 a	165 c
	179 q
	143 t
BASE COUNT	

Query Match 3.2% Score 38; DB 12; Length 632;
Best Local Similarity 55.2%; Pred. No. 12;
Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

384 TTCTTGTGATGACATCAGGATCCCGCGTCTCTACTGCGGATACGGCGCGGAG 443
190 TCCTTGGGAATATATACGATAGTATCGCTCCGCGGAGCAGCAGCAACCGGAATCCAGAA 131
444 ACTGACGCTTCAGCAGTACCATACCAAGCGTTTCATTTTCCGAAGCGATGACCAACCA 503
130 GCCATGACTCCAGCAGTACTTGTCCGAGGGTCTCAGACTCTGANGGCAATGCAATACG 71
504 CTGGCAATCCGGTA 517
70 TCCCACTGGCGTA 57

RESULT 19
BJ292385/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
BASE COUNT
ORIGIN

BJ292385 479 bp mRNA linear EST 09-APR-2002
aestivum unpublished cDNA library, wh_SL Triticum
aestivum cDNA clone whsl27102 5', mRNA sequence.
BJ292385.1 GI:20108645
EST.
bread wheat.
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 479)
Ogihara, Y. and Mural, K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadasu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. 479
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4563"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_SL"
/tissue_type="seed DPA30"
/dev_stage="Feekes' scale 11.3"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI. Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give phagescript phagemids in the TJ Clonase lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."
116 a 124 c 147 g 90 t

Query Match 3.2%; Score 37.8; DB 13; Length 479;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 81; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

371 CGTCCCTGTCACATTTCTTGTGATGACATCAGGATCCGCGGCTCTACTGCGGATAC 430
204 CTTGCGCTCTCTGCTTCTGCTGATATATATCAGTATGCTCCGCGGAGCAGCCACAC 145
431 GGCACGCGGAGACTGAGCTTTCAGCCAGTACCATACCAACAGCTTTCATTTTCCGAAGG 490
144 CGGACTCCAGAGGATGAGTCCAGGACTACTTGTCCGAGCGTCTCAGACTCTGAAGG 85

491 CATGACCAACACACTGCGCAATCCGCTAGACCGG 523
84 CATTCGAATAGTCCCTCCACTGGCATAGCGCTG 52

C20457 545 bp mRNA linear EST 03-APR-2002
C20457 Rice panicle at ripening stage Oryza sativa (japonica
cultivar-group) cDNA clone EL2053, mRNA sequence.
C20457
C20457.2 GI:11065616
EST.
Oryza sativa (japonica cultivar-group).
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 545)
Sasaki, T.
Rice cDNA from panicle at ripening stage
Unpublished (1996)
On Oct. 24, 1996 this sequence version replaced gi:1632728.
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = "RGP".
Location/Qualifiers
1. 545
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="EL2053"
/clone_lib="Rice panicle at ripening stage"
/dev_stage="ripening stage"
/note="Organ: panicle; Rice cDNA from panicle at ripening
stage"
142 a 125 c 162 g 116 t

Query Match 3.2%; Score 37.6; DB 14; Length 545;
Best Local Similarity 52.6%; Pred. No. 15;
Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

371 CGTCCCTGTCACATTTCTTGTGATGACATCAGGATCCGCGGCTCTACTGCGGATAC 430
229 CTTCCCTCTCTGATCTTCTGGAATATATCAGGATACACACAGCAGCTCGAAGGAC 170
431 GGCACGCGGAGACTGAGCTTTCAGCCAGTACCATACCAACAGCTTTCATTTTCCGAAGG 490
169 CGGACTCCAGATGACATGAGTCCAGGACTACTTGACCAAGTGTTCGAGACTCCGAGG 110
491 CATGACCAACACACTGCGCAATCCGCTAGACCGGTA 526
109 CATCAGAAACATCACCGCTGGCATATGCTGTGA 74

RESULT 21
CNS02D50/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CNS02D50 756 bp DNA linear GSS 13-MAY-2000
Tetraodon nigroviridis genome survey sequence PUC-Orl end of clone
256P09 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL192027.1 GI:7830131
GSS: genome survey sequence.
Tetraodon nigroviridis.
Tetraodon nigroviridis
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE AUTHORS

1 (bases 1 to 756)
Roest-Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL REFERENCE AUTHORS

2 (bases 1 to 756)
Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL REFERENCE AUTHORS

3 (bases 1 to 756)
Direct Submission
Submitted (12-APR-2000)

JOURNAL COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES source

1..756
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="256P09"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG256CH05SP1-end : PUC-ori"

BASE COUNT 217 a 156 c 204 g 174 t 5 others
ORIGIN

Query Match 3.2%; Score 37.4; DB 17; Length 756;
Best Local Similarity 48.8%; Pred. No. 17;
Matches 101; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 967 CAATAGTGAAGGGTGCACATATCAGACGGCTCAGTCTGTGTATATTATTCATGTC 1026
DB 333 CATTATTGAAGTGGCTGTGACCTCTTGTAGTTCATCAATCTTATTACCTGTGGG 274
QY 1027 CACTATFGCAGATACCATCAGTGTAAATTCGCCGATAATCCGTGAAGTCTGAGGA 1086
DB 273 AAATATGGCTGGCAATCACCGTACAAAGTAATGATACATAAACCATACAGAGAT 214
QY 1087 TGGAGGAGGTGAAGCTGTTCCTGAAGGATATAAGTACATCATGCCCTTTTTC 1146
DB 213 GATTCAACAGACACCACCATGCGACCTGGCAGGAAGGTAATATTTCCGCTGTTTTC 154
QY 1147 TGGCTCCGGAGCAATTTACTTTTTT 1173
DB 153 CACCTACATCAGACATCTATTGT 127

RESULT 22 AV784670 LOCUS

AV784670 RAFL6 Arabidopsis thaliana cDNA clone RAFL06-07-E03 3',
DEFINITION
mRNA sequence.

ACCESSION VERSION AV784670 KEYWORDS SOURCE

AV784670.1 GI:19803460
EST.
thale cress.

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE AUTHORS

1 (bases 1 to 649)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,

TITLE JOURNAL COMMENT

Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekieric.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES source

1..649
Location/Qualifiers
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFL06-07-E03"
/clone_lib="RAFL6"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH108"
/note="Site_1: SstI; Site_2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

BASE COUNT 182 a 173 c 102 g 192 t
ORIGIN

Query Match 3.1%; Score 37.2; DB 10; Length 649;
Best Local Similarity 52.6%; Pred. No. 20;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 373 TCCTGTGCATCTTCTGATGACATCAGGATCCGGCCGCTCTCCTGCGCATACGG 432
DB 453 TTCCCTCCCTGATCTTCAGGGATATATCAGGGATTCACGCGCAGCGCGCAGCAG 512
QY 433 GCACGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTTCGAAGGCA 492
DB 513 GAAGTCTGAGACATGCTTCAAGAACCAAGGCGAAGTCTCCGACTCTGATGGCA 572
QY 493 TGACCAACACATCGCAATCCGCTAGACCGGTAA 526
DB 573 TCACAAACACATCTCCACTTGCCTAAGCTTGTC 606

RESULT 23 BG840188 LOCUS

BG840188 MEST8-G12.T3 ISUM3-TL Zea mays cDNA clone MEST8-G12 3', mRNA
DEFINITION
sequence.

ACCESSION VERSION BG840188 KEYWORDS SOURCE

BG840188.2 GI:14243001
EST.
Zea mays.
Zea mays.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE AUTHORS

1 (bases 1 to 501)
Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks

JOURNAL COMMENT

Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14205510.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299

Email: schnable@iastate.edu

PCR Primers

FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)

BACKWARD: T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

Location/Qualifiers

1. 501
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST8-G12"
/clone_lib="ISUM3-TL"
/tissue_type="Seedling and silk"
/lab_host="DH10B"

/note="vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'
AATCGGAAGAAATTCGCCGCCGACAGGATTTTTTTTTTTTTT). The
resulting DNA-RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT73PAC vector."

114 a 135 c 105 g 147 t

FEATURES

SOURCE

FEATURES

SOURCE

Location/Qualifiers

1. 540
/organism="Sorghum propinquum"
/db_xref="taxon:132711"
/note="Organ: Floral-Induced Meristem 1 (FMI)."
/clone_lib="Floral-Induced Meristems; Vector:
paluescript II from lambda Zap II; Site_1: XhoI; Site_2:
EcoRI; mature plants were placed in a growth chamber for
15 days with 16 hr darkness and 8 hr light (flowering is
induced by short-day conditions); 16 days after being
returned to the greenhouse under natural long days during
late April/early May, meristems were harvested. The
library was made from poly-A RNA in the cloning vector
lambda ZAP II. Clones to be sequenced were prepared by
mass excision."

BASE COUNT 151 a 114 c 137 g 138 t

ORIGIN

Query Match 3.1%; Score 36.6; DB 12; Length 540;
Best Local Similarity 56.1%; Pred. No. 28;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 371 CGTCCCTGTACATTTCTTCTGATGACATCAGGATCCGCCGCTCTCACTGGCGATAC 430
Db 124 CTTACCCCTCTGCTGCTTGGGTATTATATCAGGTATCCCTCCAGCAGGAGCGACAAC 65
QY 431 GGGCAGCGCGGAGACTGAGCTTCCAGCAGTACCATACCAACAGCTTCAATTTCCGAGG 490
Db 64 TGGGACTCCAGAGCCATGACTCCAGCACTACTTTCGCCGAGGGTCTCAGACTCTGAAGG 5
QY 491 CAT 493
Db 4 CAT 2

RESULT 25

Bi968888

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Bi968888 672 bp mRNA linear EST 23-OCT-2001
GM830006B20A01 Gm-r1083 Glycine max cDNA clone Gm-r1083-1946 3',
mRNA sequence.

ACCESSION BI968888.1 GI:16343293

VERSION BI968888

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

REFERENCE 1 (bases 1 to 672)

AUTHORS Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,

Erpelting,J., Rapp,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.

A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: AW203245 corresponding to Gm-cl028-1235 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu

This clone is available through: Incyte Genomics, 4633 World

Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or

(314) 427-3222 FAX: (314) 427-3324. Web site:

http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio

n/index

Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.

Location/Qualifiers

1. 672

/organism="Glycine max"

FEATURES

source

RESULT 24

BF586903/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF586903 540 bp mRNA linear EST 12-DEC-2000
FMI_31_H07_g1_A003 Floral-Induced Meristem 1 (FMI) Sorghum
propinquum cDNA, mRNA sequence.

ACCESSION BF586903

VERSION BF586903.1 GI:11679227

KEYWORDS EST.

SOURCE Sorghum propinquum.

ORGANISM Sorghum propinquum

REFERENCE 1 (bases 1 to 540)

AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M., Marsala,C. and Pratt

,L.H.

An EST database from Sorghum: floral-induced meristems

Unpublished (2000)

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA. vector and regions

below Phred quality 16. The threshold for highest quality sequence

is 20.

Seq primer: PolyTMix

High quality sequence start: 9

High quality sequence stop: 529

POLYA-NO.

```
/db_xref="taxon:3847"
/clone="Gm-r1083-1946"
/clone_lib="Gm-r1083"
```

/note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with *Bradyrhizobium japonicum*, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.ahc.umn.edu/biodata/nfssoy/>. . . Reracking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.lie.uic.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the "OTHER EST" field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under "OTHER EST"."

BASE COUNT	167 a	166 c	140 g	174 t	25 others
ORIGIN					

Query Match	3.18;	Score 36.4;	DB 13;	Length 672;
Best Local Similarity	53.38;	Pred. No. 33;		
Matches 73;	Conservative	0;	Mismatches 64;	Indels 0

[illegible]

RESULT 26	BI498039/c	BI498039	499 bp	mRNA	linear	EST 29-NOV-2001
LOCUS		sag1sc08.y1	Gm-c1080	Glycine max	cDNA clone	GENOME SYSTEMS CLONE
DEFINITION		ID: Gm-c1080-1263	5', similar to	TR:Q9R6U1	Q9R6U1	SQDX. ; mRNA sequence.

ACCESSION	BI498039	sequence.
VERSION	BI498039.1	GT:15337383

VERSION	BI43003
KEYWORDS	EST.
SOURCE	soybean.
ORGANISM	Glycine

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE
1 (bases 1 to 499)
glycine.

1 (PAGES 1 TO 433)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Role of the Parent in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Impact of the School Environment on Learning	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the Student in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of the Teacher's Attitude on Learning	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Role of the Curriculum in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Impact of the School's Resources on Learning	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School's Culture on Learning	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of the School's Policies on Learning	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 21
South Memorial Parkway Huntsville, AL 35801 For further info
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 327.

FEATURES

source

```

1. 499
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS
/clone_lib="Gm-cl080"
/tissue_type="Roots of
mutant NTS382 seedlings
/dev_stage="8 days old"

```

/note=vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of 8 day old 'Bragg' supermodulating mutant NT5382 seedlings that were infected with *Bradyrhizobium japonicum*, strain USDA 110, 72 hours prior to harvest. Dr. Gary Stacey generously donated the tissue. The roots were flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCmP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGACATGCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II SK+ predigested vector (pBluescript II SK(+), vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Corryell at Northern Arizona University."

89 c 137 q 130 t
a

**BASE COUNT
ORIGIN**

Query Match 3.1%; Score 36.2; DB 13; Length 499;
Best Local Similarity 54.0%; Pred. No. 36;
Matches 74: Conservative 0; Mismatches 63; Indels 0; Gaps 0;

391	GAATGACATCAGGGATCCCGCCGTC	CTACTGGCGATAAAGGGCAGCGCGGAGACTGACG	450
Qy			
368	GGATATATCCGAATACCTCCACGACG	TCCGCCACCAAGCATCTCCCTGAAGACATCG	309
Db			
451	CTTCAGCCAGTACCATACCAAGCGTT	TCATTTCCGAAGGCATGACCACCACACTGGCAA	510
Qy			
308	CTTCCAAAACGACAAACCAAGCGTCT	CTGACTCTGAAGGCATGACAAATACATCTCCAC	249
Db			
511	TCCGGTAGACCGGTAAC	527	
Qy			

[illegible]

```

ACCESSION      BQ878757
VERSION        BQ878757.1
KEYWORDS       GI:22270765
SOURCE         EST.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 974)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: gcraps-remail.nih.gov
               Tissue Procurement: Dr. James R. Lupski
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Agencourt Bioscience Corporation
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLML13560 row: e column: 04
               High quality sequence start: 62
               High quality sequence stop: 560.
               Location/Qualifiers
FEATURES             source
                     1..974
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6179307"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /note="vector: PCMV-SPORT6 (Life Technologies); Site_1:
                     NOT1; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGCCG-3' and
                     5'-GACTAGTCTTAGATCGGACGCGGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      228 a   246 c   256 g   244 t
ORIGIN
Query Match      3.08; Score 35.8; DB 14; Length 974;
Best Local Similarity 65.8; Pred. No. 51;
Matches 52; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 525 AAGCTGGGAAAGGGACCTGGCATTAACACATCTCCGCTCATTCGCCAGGTGTCTGTC 584
    || || || || || || || || || || || || || || || || || || || ||
Db 478 AATCTTTAGACGAGGACCTTCGAGGAACACAGCTCCCCCACTACTGTGGACCTTC 419
    || || || || || || || || || || || || || || || || || || || ||
Qy 585 TGCTGACGCAGACGTGCTT 603
    ||| ||| ||| ||| |||
Db 418 CTCTGGGGAAGACTGGCTT 400

RESULT 31
CONS0171M/c
LOCUS
DEFINITION      Drosophila melanogaster genome survey sequence T7 end of BAC
                 BACN17E01 of DrosBAC library from Drosophila melanogaster (fruit
                 fly), genomic survey sequence.
ACCESSION      AL108088
VERSION        AL108088.1
KEYWORDS       GI:5628392
SOURCE         GSS.
ORGANISM       Drosophila melanogaster.
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.

```

[illegible]

RESULT 39
AW459606/c
LOCUS
DEFINITION
AW459606 linear
sh89c12.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl016-7343.y1 similar to TR:024302 O24302 SUCROSE SYNTHASE ;,
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AW459606
AW459606.1 GI:7029823
EST.
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1. (bases 1 to 446)
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corvelli, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1002 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 409.
FEATURES
location
1..446

```

FEATURES
  source
    align quality sequence stop: 409.
    Location/Qualifiers
      1. .446
        /organism="Glycine max"
        /db_xref="taxon:3847"
        /clone="GENOME SYSTEMS CLONE ID: Gm-cl016-7343"
        /clone_lib="Gm-cl016"
        /tissue_type="immature flowers of field grown plants"
        /lab_host="Xl10-Gold"
        /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
        /XhoI; This cDNA library was constructed from mRNA isolated
        from immature flowers of field grown plants. The cDNA
        library was prepared using the Stratagene pBluescript II
        XR library construction kit. Complementary DNA was
        synthesized from mRNA using a primer consisting of a poly
        (dT) sequence with a XhoI restriction site. EcoRI adapters
        were ligated to the blunt-ended cDNA fragments followed by
        XhoI digestion. The cDNA fragments were directionally
        cloned into the EcoRI-XhoI restriction site of the
        pBluescript vector. The ligated cDNA fragments were
        transformed into Xl10-Gold host cells. This library was
        constructed by Dr. Randy Shoemaker and Dr. John
        Erpelding."
      BASE COUNT      108 a      99 c      133 g      106 t
      ORIGIN
        Query Match      2.98; Score 34.8; DB 10; Length 446;
        Best Local Similarity 55.9%; Pred. No. 89;
        Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
      Qy      387 TTCTGTAATGACATCAGGGAATCCGCGCGTCTCACTGGCGATAACGGGCACGCCGAGACT 446
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      Db      334 TGCACATGATCTCAGCAGGACCACCATTGCATCTGGCGAATGTGGCAAGCCGCAAGTC 275
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      Qy      447 GAGCCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAAGCATGACCAACAC 504
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      Db      274 ATGCCCTCAACCATGTGCAACAAAGCCCTCGTATACAGCAGCGTGCACGAAGCAC 217
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 40
 A1496532/c
 LOCUS
 DEFINITION
 A1496532 478 bp mRNA linear EST 30-NOV-2001
 sb13a03.y1 Gm-cl004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl004-8645 5' similar to SW:SUSY_PHAU Q01390 SUCROSE SYNTHASE
 //, mRNA sequence.
 ACCESSION
 VERSION A1496532.1 GI:4397535
 KEYWORDS
 SOURCE soybean.
 ORGANISM
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 478)
 Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Willson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand This clone is
 available through: ResGen, Invitrogen Corp. 2130 South Memorial
 Parkway Huntsville, AL 35801 For further information call: (800
)-533-4363 or contact via email: ccu@resgen.com
 Seq primer: -40RP from Gbco
 High quality sequence stop: 1
 POLYA-No.
 Location/Qualifiers
 1..478
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl004-8645"
 /clone.lib="Gm-cl004"
 /tissue.type="root"
 /lab_host="XL10-Gold"
 /note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2:
 XhoI; Root cDNA. The mRNA was isolated from entire roots
 of 8 day old 'Williams' seedlings which were propagated on
 paper towels with distilled water. Stratagene's cDNA
 Synthesis Kit (catalog #200401) was used to synthesize the
 cDNA. First- strand synthesis was performed with 5-methyl
 dCTP, hence the ligated cDNA is hemimethylated.
 Stratagene's first-strand synthesis primer was used
 [CAGAGAGAGAGAGAGAGAGTAGTCTGAG(T)-18]. After
 second-strand synthesis, the cDNA ends were 'polished'
 with clone pfu DNA polymerase, ligated to EcoRI adapters,
 and phosphorylated. The XhoI site within the first-strand
 synthesis primer was restricted by digestion with XhoI;
 all XhoI sites in the cDNA would be protected by their
 hemimethylated status. The cDNA constructs were
 size-fractionated with a 500bp cutoff, using GbcoBRL Life
 Technologies' cDNA Size Fractionation column. The column
 eluent was then ligated into Stratagene's pBluescript II
 XR predigested vector (pBluescript II SK(+)) that had been
 digested with EcoRI and XhoI, and phosphorylated). Both
 the white and blue colonies appear to contain recombinant
 plasmids with cDNA inserts. Blue colonies 9p-15) have been
 sequenced, and possess putative cDNA inserts. This library
 was constructed by Dr. Paul Keim & Virginia H. Coryell,

Department of Biology, Box5640, Northern Arizona
 University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
 Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:
 520-523-7500, email: paul.keim@na.u.edu,
 virginia.coryell@na.u.edu"

BASE COUNT 115 a 98 c 148 g 115 t 2 Others
 ORIGIN

Query Match 2.9%; Score 34.8; DB 9; Length 478;
 Best Local Similarity 55.9%; Pred. No. 90;
 Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCAGGATCCGCCCTCTCTCCTGCGTAAACGGGACCGGAGACT 446
 Db 342 TGCACAAATGATCTCAGCAGCACCACTTGTGGGCGATGTTGGCAAGCCGCAAGTC 283
 QY 447 GAGCGTTTCAGCCAGTACCATACCAACCGTTTCATTTCCGAGGAGTCACCAACACAC 504
 Db 282 ATGGCGCTCAACCCATGCTCAACCAAAAGCCTGATATACAGCAGGCTGCACGAAGCAC 225

RESULT 41
 AW620859/c
 LOCUS
 DEFINITION
 AW620859 507 bp mRNA linear EST 03-DEC-2001
 sj47ell.y1 Gm-cl033 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl033-453 5' similar to SW:SUSY_SOYBN P13708 SUCROSE SYNTHASE ;,
 mRNA sequence.
 ACCESSION
 VERSION AW620859.1 GI:7327061
 KEYWORDS
 SOURCE soybean.
 ORGANISM
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 507)
 Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Willson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Insert Length: 838 Std Error: 0.00
 High quality sequence stop: 419.
 Location/Qualifiers
 1..507
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl033-453"
 /clone.lib="Gm-cl033"
 /tissue.type="Desloy 5710' seedling roots"
 /lab_host="DH108"

note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from 'Desloy 5710' seedling roots. Tissue was taken from
 7-day-old seedlings that had been propagated on paper
 towels with distilled water. Tissue was taken from the tip
 to the first lateral root, usually about 3cm from the tip,
 and flash-frozen in liquid nitrogen. Stratagene's cDNA
 Synthesis Kit (catalog number 200401) was used to

seq primer: T/

```
/clone="Plate=1031 Col=4 Row=1"
```

```
/clone="Plate=1031 Col=4 Row=1"
```



```

FEATURES
  source
    Location/Qualifiers
      1..540
        /organism="Glycine max"
        /db_xref="taxon:3847"
        /clone="GENOME S1SYSTEMS CLONE ID: Gm-cl086-1004"
        /clone_lib="Gm-cl086"
        /tissue_type="young seeds (Williams 82)"
        /lab_host="DH10B"
        /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:

```

```

FEATURES
  source
    Location/Qualifiers
      1..542
        /organism="Glycine max"
        /db_xref="taxon:3847"
        /clone="GENOME SYSTEMS CLONE ID: Gm-cl049-4322"
        /clone_lib="Gm-cl049"
        /tissue_type="whole seedlings of greenhouse grown plants"

```

```

/dev_stage="3 week old"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The Clark NIL was constructed and seed was provided
by Dr. J. Specht, University of Nebraska (Shoemaker and
Specht, 1995). The cDNA library was constructed from mRNA
isolated from whole seedlings of 3 week old greenhouse
grown plants. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (GibcoBRL). The library was constructed
in cooperation with Dr. Paul Keim's laboratory at Northern
Arizona University."
BASE COUNT      118 a      128 c      139 g      157 t
ORIGIN

```

```

Query Match      2.9%; Score 34.8; DB 13; Length 542;
Best Local Similarity 55.9%; Pred. No. 91;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCAGGATCCGCCGCTCTACTGCGATACACGGCGGAGACT 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 TGCACATGATCTCAGCAGGACCAACCATTCATGTGCGGAATGTTGGCAAGCGCAAGTC 78
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 447 GACGCTTCAGCAGTACCAACCAACGCTTCATTTCCGAAGGATGACCAACAC 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ATGGCCTCAACCACTGTCAAAACCAAAAGCCTGTATACAGCAGGTCGCAAGAACAC 20
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 48
LOCUS BI427022/c
DEFINITION sgi0403.y1 Gm-cl080 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl080-558 5' similar to SW:SUSY_SOYBN P13708 SUCROSE
SYNTHASE ;, mRNA sequence.
ACCESSION BI427022.1 GI:15204254
VERSION EST.
KEYWORDS soybean.
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

```

```

REFERENCE 1 (bases 1 to 566)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 391.
Location/Qualifiers
1..566
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl080-558"

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source

```

```

/clone_lib="Gm-cl080"
/tissue_type="Roots of 8 day old 'Bragg' supernodulating
mutant NTS382 seedlings"
/dev_stage="8 days old"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The mRNA was isolated from roots of 8 day old
'Bragg' supernodulating mutant NTS382 seedlings that were
infected with Bradyrhizobium japonicum, strain USDA 110, 72
hours prior to harvest. Dr. Gary Stacey generously donated
the tissue. The roots were flash-frozen in liquid
nitrogen. StrataGene's cDNA Synthesis Kit (catalog number
200401) was used to synthesize the cDNA. First-strand
synthesis was performed with 5-methyl dCTP, hence the
ligated cDNA was hemimethylated. A modification of
StrataGene's first-strand synthesis primer was used. An
'anchor' nucleotide (V=A, C, or G) was added to the 3' end
of the primer [GAGAGAGAGAGAGAGAGACTGCTCGAG(T18V)] to
anchor the primer at the 5' end of the poly(A) tract.
After second-strand synthesis, the cDNA ends were filled
in with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The cDNA was then
precipitated and redissolved in sterile, RNase-
DNase-free water. The XhoI site within the first- strand
synthesis primer was then restricted by digestion with
XhoI from Promega (400/ul); all XhoI sites in the cDNA
would be protected by their hemimethylated status. The
cDNA constructs were size-fractionated with a 500bp
cutoff, using Sephacryl S-500 High Resolution (Pharmacia
Biotech) in a 2-mm diameter column and a bed volume of
approximately 1ml. The column eluent was precipitated,
redissolved, and ligated into StrataGene's pBluescript II
XR Predigested vector (pBluescript II SK(+)) vector that
has been digested with EcoRI and XhoI, and phosphorylated
by StrataGene). This library was constructed in the
laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at
Northern Arizona University."
BASE COUNT      136 a      135 c      155 g      140 t
ORIGIN

```

```

Query Match      2.9%; Score 34.8; DB 13; Length 566;
Best Local Similarity 55.9%; Pred. No. 92;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCAGGATCCGCCGCTCTACTGCGATACACGGCGGAGACT 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 TGCACATGATCTCAGCAGGACCAACCATTCATGTGCGGAATGTTGGCAAGCGCAAGTC 224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 447 GACGCTTCAGCAGTACCAACCAACGCTTCATTTCCGAAGGATGACCAACAC 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 ATGGCCTCAACCACTGTCAAAACCAAAAGCCTGTATACAGCAGGTCGCAAGAACAC 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 49
LOCUS BQ630772/c
DEFINITION BQ630772.1 Gm-cl082 Glycine max cDNA clone SOYBEAN CLONE ID:
ID: Gm-cl082-3884 5' similar to SW:SUSY_SOYBN P13708 SUCROSE SYNTHASE
SYNTHASE ;, mRNA sequence.
ACCESSION BQ630772
VERSION EST.
KEYWORDS soybean.
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

```

```

REFERENCE 1 (bases 1 to 581)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 391.
Location/Qualifiers
1..566
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl080-558"

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source

```

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Socioeconomic Status on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Education	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Training on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Socioeconomic Status on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

REFERENCE
AUTHORS

TITLE	AUTHORS
JOURNAL	
COMMENT	

FEATURES

B1761211
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMS
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
sour

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:08:24 ; Search time 46.8861 Seconds
(without alignments)
7724.783 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcaggagatgaaataaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.2	4.8	4403765	4	US-09-103-840A-2
2	57.2	4.8	4411529	4	US-09-103-840A-1
3	39.8	3.4	4403765	4	US-09-103-840A-2
4	39.8	3.4	4411529	4	US-09-103-840A-1
5	33.8	2.9	1647	6	5405943-3
6	33	2.8	3147	4	US-09-066-046-1
7	33	2.8	5570	4	US-09-066-046-7
8	32.6	2.8	500	3	US-09-141-000-2
9	32.6	2.8	2426	3	US-09-028-327-2
10	32.2	2.7	3519	1	US-08-035-558-1
11	32.2	2.7	7218	1	US-08-232-463-14
12	31.8	2.7	7218	1	US-08-232-463-14
13	31.4	2.7	80161	3	US-09-036-987A-1
14	31.4	2.7	80161	4	US-09-370-700-1
15	31.2	2.6	1425	1	US-08-353-550-7
16	31.2	2.6	1425	2	US-08-551-687-7
17	31.2	2.6	3031	4	US-08-664-962B-3
18	31.2	2.6	3031	4	US-09-311-743-3
19	31.2	2.6	3259	5	PCT-US95-03747-1
20	31.2	2.6	8802	3	US-08-896-449A-1
21	31.2	2.6	8802	3	US-09-132-652-1
22	30.8	2.6	1122	4	US-09-403-768-7
23	30.8	2.6	6474	4	US-08-961-527-155
24	30.6	2.6	1399	1	US-07-737-736B-5
25	30.4	2.6	780	4	US-09-651-656-26
26	30.4	2.6	780	4	US-09-650-855-26
27	30.4	2.6	3382	2	US-08-682-847-1

28	30.2	2.6	387	4	US-08-765-907A-4	Sequence 4, Appl
29	30.2	2.6	2888	4	US-08-765-907A-1	Sequence 1, Appl
30	29.6	2.5	4608	4	US-09-041-886-24	Sequence 24, Appl
31	29.6	2.5	4608	5	PCT-US94-05277-1	Sequence 1, Appl
32	29.4	2.5	1134	4	US-09-482-180A-3	Sequence 3, Appl
33	29.2	2.5	2380	1	US-08-572-951-3	Sequence 3, Appl
34	29.2	2.5	3865	1	US-08-832-883-48	Sequence 48, Appl
35	29.2	2.5	3865	2	US-08-832-877-48	Sequence 8, Appl
36	29.2	2.5	4004	4	US-09-293-505-8	Sequence 5, Appl
37	29	2.5	748	6	5405943-1	Patent No. 5405943
38	29	2.5	1426	4	US-08-858-207A-78	Sequence 78, Appl
39	29	2.5	5338	4	US-08-961-527-50	Sequence 50, Appl
40	29	2.5	25165	4	US-09-453-702B-39	Sequence 39, Appl
41	28.6	2.4	668	3	US-08-578-634C-2	Sequence 2, Appl
42	28.6	2.4	668	4	US-09-430-010-2	Sequence 2, Appl
43	28.6	2.4	1730	4	US-09-575-574-5	Sequence 5, Appl
44	28.6	2.4	2563	1	US-08-608-241-5	Sequence 5, Appl
45	28.6	2.4	2563	2	US-08-922-182-5	Sequence 5, Appl
46	28.6	2.4	2563	2	US-08-919-953-5	Sequence 5, Appl
47	28.6	2.4	3572	4	US-09-575-574-3	Sequence 3, Appl
48	28.4	2.4	621	4	US-09-180-109A-17	Sequence 17, Appl
49	28.4	2.4	1002	1	US-08-279-058B-13	Sequence 13, Appl
50	28.4	2.4	1002	4	US-08-828-323-13	Sequence 13, Appl
51	28.4	2.4	1544	2	US-08-837-593-1	Sequence 1, Appl
52	28.4	2.4	2940	4	US-09-703-416-1	Sequence 1, Appl
53	28.4	2.4	4608	1	US-07-718-575-11	Sequence 11, Appl
54	28.4	2.4	4608	1	US-08-481-206-11	Sequence 11, Appl
55	28.4	2.4	4608	2	US-08-486-269A-11	Sequence 11, Appl
56	28.4	2.4	4700	4	US-09-150-460B-9	Sequence 9, Appl
57	28.2	2.4	622	4	US-09-385-982-46	Sequence 46, Appl
58	28.2	2.4	805	4	US-08-964-722-1	Sequence 1, Appl
59	28.2	2.4	1373	4	US-09-257-179-33	Sequence 33, Appl
60	28.2	2.4	1470	1	US-08-599-252-97	Sequence 97, Appl
61	28.2	2.4	1470	5	PCT-US96-06352-97	Sequence 97, Appl
62	28.2	2.4	1470	5	PCT-US96-06583-97	Sequence 97, Appl
63	28.2	2.4	3633	4	US-09-221-017B-362	Sequence 362, App
64	28.2	2.4	4899	2	US-08-210-535-5	Sequence 5, Appl
65	28.2	2.4	6415	4	US-08-961-527-162	Sequence 162, App
66	28.2	2.4	38564	4	US-09-734-673-3	Sequence 3, Appl
67	28	2.4	1001	4	US-09-641-638-554	Sequence 554, App
68	28	2.4	5407	4	US-09-453-702B-201	Sequence 201, App
69	28	2.4	8625	4	US-08-980-832-1	Sequence 1, Appl
70	28	2.4	11233	4	US-08-980-832-27	Sequence 27, Appl
71	27.8	2.4	466	4	US-09-141-027-5	Sequence 5, Appl
72	27.8	2.4	498	6	5514567-3	Patent No. 5514567
73	27.8	2.4	501	4	US-09-331-260-1	Sequence 1, Appl
74	27.8	2.4	561	1	US-08-213-448-2	Sequence 2, Appl
75	27.8	2.4	561	3	US-08-912-768-2	Sequence 2, Appl
76	27.8	2.4	561	5	PCT-US95-03206-2	Sequence 2, Appl
77	27.8	2.4	771	2	US-08-972-008-3	Sequence 3, Appl
78	27.8	2.4	771	4	US-09-267-409-3	Sequence 3, Appl
79	27.8	2.4	772	6	5326859-2	Patent No. 5326859
80	27.8	2.4	777	6	5514567-2	Patent No. 5514567
81	27.8	2.4	1142	1	US-08-006-082A-5	Sequence 5, Appl
82	27.8	2.4	1335	4	US-09-859-698-1	Sequence 1, Appl
83	27.8	2.4	1335	4	US-09-428-979-1	Sequence 1, Appl
84	27.8	2.4	1437	4	US-09-137-223A-4	Sequence 4, Appl
85	27.8	2.4	1584	4	US-08-406-030A-24	Sequence 24, Appl
86	27.8	2.4	2495	4	US-09-141-027-1	Sequence 1, Appl
87	27.8	2.4	2525	2	US-08-972-008-1	Sequence 1, Appl
88	27.8	2.4	2525	4	US-09-267-409-1	Sequence 1, Appl
89	27.8	2.4	3109	3	US-08-938-291A-1	Sequence 1, Appl
90	27.8	2.4	4268	1	US-08-264-534-33	Sequence 33, Appl
91	27.8	2.4	4268	1	US-08-083-590A-12	Sequence 12, Appl
92	27.8	2.4	4268	1	US-08-465-500-33	Sequence 33, Appl
93	27.8	2.4	4268	2	US-08-346-128-33	Sequence 33, Appl
94	27.8	2.4	4268	3	US-08-532-384-12	Sequence 12, Appl
95	27.8	2.4	4268	3	US-08-893-828-33	Sequence 33, Appl
96	27.8	2.4	6701	4	US-09-004-838-107	Sequence 107, App
97	27.8	2.4	9723	1	US-08-083-590A-21	Sequence 21, Appl
98	27.8	2.4	9723	2	US-08-532-384-21	Sequence 21, Appl
99	27.8	2.4	10747	2	US-08-147-777-1	Sequence 1, Appl
100	27.8	2.4	10747	3	US-08-452-872-1	Sequence 1, Appl

c 101 27.8 2.4 10747 5 PCT-US93-03985-1
c 102 27.8 2.4 13149 4 US-09-004-838-87
c 103 27.8 2.4 15062 3 US-09-004-838-89
c 104 27.8 2.4 15202 3 US-08-922-635-21
c 105 27.8 2.4 24979 2 US-08-147-777-3
c 106 27.8 2.4 24979 3 US-08-452-872-3
c 107 27.8 2.4 24979 5 PCT-US93-03985-3
c 108 27.8 2.4 50341 1 US-08-247-901C-1
c 109 27.8 2.4 50341 2 US-09-075-904-1
c 110 27.8 2.4 52297 4 US-09-426-436-1
c 111 27.8 2.4 52297 4 US-08-705-557-1
c 112 27.6 2.3 215 1 US-08-238-163-5
c 113 27.6 2.3 227 4 US-09-397-787-119
c 114 27.6 2.3 1001 4 US-09-641-638-553
c 115 27.6 2.3 2032 4 US-09-500-063-1
c 116 27.6 2.3 2235 4 US-09-334-601-3
c 117 27.6 2.3 2288 4 US-09-334-601-1
c 118 27.6 2.3 3243 1 US-08-611-107-32
c 119 27.6 2.3 3494 4 US-09-334-601-5
c 120 27.4 2.3 218 1 US-08-248-474-100
c 121 27.4 2.3 218 3 US-08-756-849-100
c 122 27.4 2.3 498 4 US-09-328-111-252
c 123 27.4 2.3 540 4 US-08-793-109-13
c 124 27.4 2.3 600 2 US-08-679-217-1
c 125 27.4 2.3 804 4 US-09-134-001C-2575
c 126 27.4 2.3 925 4 US-09-221-017B-927
c 127 27.4 2.3 2219 3 US-08-510-646B-17
c 128 27.4 2.3 162450 4 US-09-345-882-1
c 129 27.4 2.3 1374 1 US-08-123-761A-2
c 130 27.2 2.3 1455 4 US-09-615-192A-245
c 131 27.2 2.3 1956 3 US-08-693-940-2
c 132 27.2 2.3 1956 4 US-09-566-660-2
c 133 27.2 2.3 2266 1 US-08-453-472-1
c 134 27.2 2.3 2266 1 US-08-453-952-1
c 135 27.2 2.3 2266 2 US-08-484-993B-42
c 136 27.2 2.3 2266 2 US-08-862-903-1
c 137 27.2 2.3 2266 2 US-08-484-158B-42
c 138 27.2 2.3 2266 2 US-08-484-596A-42
c 139 27.2 2.3 2266 2 US-08-480-150A-42
c 140 27.2 2.3 2266 3 US-08-458-731-42
c 141 27.2 2.3 2266 3 US-08-149-223A-42
c 142 27.2 2.3 2291 1 US-07-872-644-5
c 143 27.2 2.3 2291 1 US-08-297-494-5
c 144 27.2 2.3 2291 1 US-08-297-510-5
c 145 27.2 2.3 2291 1 US-08-479-532-5
c 146 27.2 2.3 2291 1 US-08-455-526-5
c 147 27.2 2.3 2291 1 US-08-455-525-5
c 148 27.2 2.3 2291 3 US-09-139-491-5
c 149 27.2 2.3 2291 5 PCT-US92-03222-5
c 150 27.2 2.3 2291 5

ALIGNMENTS

RESULT 1
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765

; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Query Match 4.8%; Score 57.2; DB 4; Length 4403765;
Best Local Similarity 54.2%; Pred. No. 1.le-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTGTGGCGTGGCTGCATCCGGAAGGCTTCGCGTGGGATAAAAAATCGCGCAG 331
Db 2447373 TTCTGGCGGGCGGCCCATCGGACGGCGGATCCCGATGATCAGCACTCGGCCAC 2447432
QY 332 TGGCGCGGTCCATGACAGACATCCCGGAGTACACGGTCCCTGTACATTTCTCTG 391
Db 2447433 GCGTTCGGCGACCCCGGTCCACGACCTACCGTGGACCACTAGCCAGTCTTGTGTGCG 2447492
QY 392 AATGACATCAGGATCCCGCGCTCTCACTGGCGATACGGGCACGCCGAGACTGACGC 451
Db 2447493 CACCGTTTCGGGGCTCCGCGAGAAATGCGGGGATTACCGGCACGCCGCGGAGGC 2447552
QY 452 TTCAGGACACGATGCCAAGCCCTCGACGTCC 485
Db 2447553 TTCAGGACACGATGCCAAGCCCTCGACGTCC 2447586
RESULT 2
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1
Query Match 4.8%; Score 57.2; DB 4; Length 4411529;
Best Local Similarity 54.2%; Pred. No. 1.le-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTGTGGCGTGGCTGCATCCGGAAGGCTTCGCGTGGGATAAAAAATCGCGCAG 331
Db 2450072 TTCTGGCGGGCGGCCCATCGGACGGCGGCCGATCCCGATGATCAGCACTCGGCCAC 2450131
QY 332 TGGCGCGGTCCATGACAGACATCCCGGAGGTTAACAGGGTCCCTGTACATTTCTCTG 391
Db 2450132 GCGTTCGGCCACCCCGGTCCACCGACCTACCGTTCGACCACTAGCCAGTCTTGTGTGCTG 2450191
QY 392 AATGACATCAGGATCCCGCGCTCTCACTGGCGATACGGGCACGCCGAGACTGACGC 451
Db 2450192 CACCGTTTCGGGGCTCCGCGAGAAATGCGGGGATTACCGGCACGCCGCGGAGGC 2450251
QY 452 TTCAGCCAGTACCATACCAAAACGCTTCATTTTCC 485
Db 2450252 TTCAGGACACGATGCCAAGCCCTCGACGTCC 2450285
; SEQ ID NO 2
; LENGTH: 4403765

```

RESULT 3
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          3.4%; Score 39.8; DB 4; Length 4403765;
Best Local Similarity 50.8%; Pred. No. 0.65;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 333 GCCTGGTCCATGACGACATCCCGCCGCTCTCACTGGCGATACGCGGAGACTGACGCT 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577989 GCGTGGCCCGCCACTGACGACCTCGTGGCCGACACCGAGGTGCGGTGATCCCGTCCGC 577930

Qy 393 ATGACATCAGGATCCCGCCGCTCTCACTGGCGATACGCGGAGACTGACGCT 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577929 ACCGCGAGGCGACCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 577870

Qy 453 TCAGCCAGTAGTACCAATACAAAGCTTTCATTTCCGAAGGATGACACCACTGGCAATC 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577869 TCCACAGCAACAGCGCGGACGACTCGGAGTAGTCTCGGACCGACCAAGGTCGCGGCC 577810

Qy 513 CGGTAGA 519
    ||| |||
Db 577809 CGAAACA 577803

RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          3.4%; Score 39.8; DB 4; Length 4411529;
Best Local Similarity 50.8%; Pred. No. 0.65;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```

```

Qy 333 GCCTGGTCCATGACGACATCCCGCCGCTCTCACTGGCGATACGCGGAGACTGACGCT 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576547 GCGTGGCCCGCCACTGACCGACCTCGTGGCCGACACCGAGGTGCGGTGATCCCGTCCGC 576488

Qy 393 ATGACATCAGGATCCCGCCGCTCTCACTGGCGATACGCGGAGACTGACGCT 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576487 ACCGCGAGGCGACCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 576428

Qy 453 TCAGCCAGTAGTACCAATACAAAGCTTTCATTTCCGAAGGATGACACCACTGGCAATC 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576427 TCCACAGCAACAGCGCGGACGACTCGGAGTAGTCTCGGACCGACCAAGGTCGCGGCC 576368

Qy 513 CGGTAGA 519
    ||| |||
Db 576367 CGAAACA 576361

RESULT 5
5405943-3/c
; Patent No. 5405943
; APPLICANT: COMINGS, DAVID E.
; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED
; BEHAVIORS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/562,596
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 125,577
; FILING DATE: 25-NOV-1987
; APPLICATION NUMBER: 271,653
; FILING DATE: 16-NOV-1988
; APPLICATION NUMBER: 410,831
; FILING DATE: 22-SEP-1989
; SEQ ID NO:3
; LENGTH: 1647
5405943-3

Query Match          2.9%; Score 33.8; DB 6; Length 1647;
Best Local Similarity 53.4%; Pred. No. 0.63;
Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 1013 ATTACTGTCTATGGCCACTATGGCAGATGACGAGTAAATTCCTGATTAATTCGATCCG 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 ATCTCTGACATGGCCATTCTGAAAGATCTCTCGAACAGAAATCCCACTCCGAGAGATTG 286

Qy 1073 TCGAAGTCTGAGGATGAAGGAGTGAAGGCTCTTCTTCTGAAAGGATAAAGTGACATC 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 CTTAACCAGAGTTCAATGAGTTATGATAAAAGATGTTTCATCATGATGATTTT 226

Qy 1133 ATGCCCTCTTTT 1145
    ||| ||| ||| |||
Db 225 ATTTCCTTTTGT 213

RESULT 6
US-09-066-046-1/c
; Sequence 1, Application US/09066046A
; Patent No. 6204252
; GENERAL INFORMATION:
; APPLICANT: MURPHY, Cheryl
; STOREY, James
; BELTZ, Gerald A.
; COUGHLIN, Richard T.
; TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA

```

COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,046A
FILING DATE: 24-Apr-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Superko, Colleen
REGISTRATION NUMBER: 39,850
REFERENCE/DOCKET NUMBER: 106,941,155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-6000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3147 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-066-046-1

Query Match 2.8%; Score 33; DB 4; Length 3147;
Best Local Similarity 48.2%; Pred. No. 1.7;
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 397 CATCAGGATCCGCCGCTCTCACTGGCGATACGGGCACCCGAGACCTGACGCTTCAG 456
DB 1253 CAACAGTTATCTCTCAGTAGTATCTTCGAGCAGCGTAACTGCTGTTGCGGATCTG 1194

QY 457 CCAGTACCACCAACGCTTCATTTCCGAAGGCGATGACACACACATGCGCAATCCGGT 516
DB 1193 GCACATATGATGATCTACTTCTGCTGAGTAGTCTTTTACTTGGTTAAACAACCGCAT 1134

QY 517 AGACCGGTAACTGGGAAGGCGACCTGCCATTAAACATCCTCCGCTCATTCGCCAGGT 576
DB 1133 TTGAGGCTCTGTTGCTATAGGGCGATAGACATAAGTACAGATGGTTGCACCTCCCACTG 1074

QY 577 GTTCTGCTGCTG 589
DB 1073 ATTCTTGTGCTG 1061

RESULT 7
US-09-066-046-7/c
Sequence 7, Application US/09066046A
Patent No. 6204252
GENERAL INFORMATION:
APPLICANT: MURPHY, Cheryl
STOREY, James
BELTZ, Gerald A.
COUGHLIN, Richard T.
TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,046A
FILING DATE: 24-Apr-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Superko, Colleen
REGISTRATION NUMBER: 39,850
REFERENCE/DOCKET NUMBER: 106,941,155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-6000
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5570 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-066-046-7

Query Match 2.8%; Score 33; DB 4; Length 5570;
Best Local Similarity 48.2%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 397 CATCAGGATCCGCCGCTCTCACTGGCGATACGGGCACCCGAGACCTGACGCTTCAG 456
DB 1007 CAACAGTTATCTCTCAGTAGTATCTTCGAGCAGCGTAACTGCTGTTGCGGATCTG 948

QY 457 CCAGTACCACCAACGCTTCATTTCCGAAGGCGATGACACACACATGCGCAATCCGGT 516
DB 947 GCACATATGATGATCTACTTCTGCTGAGTAGTCTTTTACTTGGTTAAACAACCGCAT 888

QY 517 AGACCGGTAACTGGGAAGGCGACCTGCCATTAAACATCCTCCGCTCATTCGCCAGGT 576
DB 887 TTGAGGCTCTGTTGCTATAGGGCGATAGACATAAGTACAGATGGTTGCACCTCCCACTG 828

QY 577 GTTCTGCTGCTG 589
DB 827 ATTCTTGTGCTG 815

RESULT 8
US-09-141-000-2
Sequence 2, Application US/09141000
Patent No. 6054295
GENERAL INFORMATION:
APPLICANT: Chen, Fang
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE REFERENCE: 19999Y
CURRENT APPLICATION NUMBER: US/09/141,000
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 500
TYPE: PRT
ORGANISM: Human
US-09-141-000-2

Query Match 2.8%; Score 32.6; DB 3; Length 500;
Best Local Similarity 10.8%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 43; Conservative 110; Mismatches 246;

QY 733 GCGGAACATTGCTGCTGATGACGCCAGGCGGTGATGACGGGTAAACGGTCCGATGG 792
DB 24 SSGIDALSHRSPSGSSDASGGFGIALGTHANGLDSPPMFAGAGLGGTTCRKSIEDCASGI 83

Qy	793	CTTCATATTGCAATGCGGGCCAGTCGGAACCCGGTGGGAATAACCGTTACCGGTGTCCTGA	852
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	84	MEDSAICEYMLNAIPKRLCLVCGDIIASGYHYGVASCEACKAFKKRTIQGNIEYSCPATN	143
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	853	CACCTTCCGCCATCAGATCGCCATCATGGTGATAGGACACAAACAAATGAATCACACA	912
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	144	ECEITRRKRKSCQACRFMKCLKVGMUKBEGVRLDVRVGRGQRYKRRLDSESPYLSLQISP	203
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	913	GATAAATTCAGGGAAAACGTTCTGCTCTTACCGGTGATGTAGGTTTTTGTCTGCACAATAG	972
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	204	PAKPLTKIVSYLLVIAEPDKLYAMPMPGMEGDIKALTLLCDLADRELVIIGWAKHIPG	263
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	973	TGAAGCGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTAATCTCATGCGCCACTAT	1032
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	264	FSSISLGDQMSLLQSAWMEILLIGIVYRSIPYDDKLVIYAEDYIMDEHSRLAGLLELYRA	323
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	1033	GGCAGATGACCATCAGATGATTTAAATTCGCCGATATCCGTGCAAGTCTTGAGGATGGAAG	1092
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	324	ILQIVRRYKRLKVEKEFVTLKALALANSMSVIEDLEAVQKLDLHLLEALQDVELSORH	383
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	1093	GAAGGTGAAGCGTGTCTCCTGAAAGGAATAAAATGACAT	1131
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	384	EEPWRTGKLLTLLPLLRQATAKAVQHIFYSVKLGQKVPMH	422
		: : : : : : : : : : : : : : : : : : : : : : : :	

RESULT 9
 US-09-028-327-2/c
 : Sequence 2, Application US/09028327
 : Patent No. 6130064
 : GENERAL INFORMATION:
 : APPLICANT: Tang, Y. Tom
 : APPLICANT: Corley, Neil C.
 : APPLICANT: Guglier, Karl J.
 : TITLE OF INVENTION: HUMAN SMN-LIKE PROTEIN
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Dr.
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/028,327
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER:
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J
 : REGISTRATION NUMBER: 36,749
 : REFERENCE/DOCKET NUMBER: PF-0482 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 650-853-0555
 : TELEFAX: 650-845-4166
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2426 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: BRSTNOT24
 : CLONE: 3769729
 : US-09-028-327-2

```

Query Match          2.8%  Score 32.6;  DB 3;  Length 2426;
Best Local Similarity 54.6%  Pred. No.2;
Matches 65;  Conservative . 0;  Mismatches 54;  Indels 0;  Gaps 0;

QY      800  ATCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTTACCGGTGTCTGTACACCTTC 859
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      254  AGCGGTTTCGGGGATCGACAGAGCCACTGTAAATATATAGACACAAGTGACTTGACCCGTC 195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      860  CGCCATCAGATCGCCATCATGGTGTAGATAGGCACAAACAAATGAAATTCACACAGATAAT 918
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      194  CTCCTTAGAACCACTATTTCAGCAAAACAGCCTATAAATTGACTATCTCTGATTAT 136
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-08-035-558-1/c
; Sequence 1, Application US/08035558
; Patent No. 5462734
; GENERAL INFORMATION:
; APPLICANT: Letchworth, Geoffrey J.
; APPLICANT: Israel, Barbara A.
; TITLE OF INVENTION: HERPESVIRUS VACCINE AND METHOD
; TITLE OF INVENTION: OF USING SAME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles and Brady
; STREET: P.O BOX 2113
; STREET: FIRST WISCONSIN PLAZA
; CITY: MADISON
; STATE: WISCONSIN
; COUNTRY: U.S.A.
; ZIP: 53701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 800Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/035,558
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/607,794
; FILING DATE: 02-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Nicholas J. Seay
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 9629691682
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608)251-5000
; TELEFAX: (608)251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: bovine herpesvirus-1
; INDIVIDUAL ISOLATE: P8-2 strain
; FEATURE:
; NAME/KEY: coding sequence for gI glycoprotein
; LOCATION: 379 to 3165
; IDENTIFICATION METHOD: sequence analysis
; PUBLICATION INFORMATION:
; AUTHORS: Misra, Vikram
; AUTHORS: Nelson, Randy
; AUTHORS: Smith, Michael
; TITLE: Sequence of a Bovine Herpesvirus Type-1 etc.
; JOURNAL: Virology
; VOLUME: 166

```

```

; PAGES: 542-549
; DATE: 1988
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 379 TO 3165
; US-08-035-558-1

Query Match
Best Local Similarity 2.7%; Score 32.2; DB 1; Length 3519;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 296 CCGAAGCGTTCCGGTCGGGATAAAATCCGCGAGTCCGCGGTCATCGACACATC 355
Db 196 CCGNGCCGCGCGGATCGGAAGCAAAAAAATGCTGAAGTCGTTACCGCCGCGCC 137
QY 356 CCCACGGGTAAACAGGTCCTGTGCATCTTCTGAATGACATCAGGGATCCGCCGT 415
Db 136 CCGCGCGCCGCGGATCGGAAGCAAAAAAATGCTGAAGTCGTTACCGCCGCGCC 77
QY 416 CTCACGCGGATAACGGCGGAGAGTCCGCGGATCGAGCTTCAGCCAGT 461
Db 76 GTTTCCAGCTAAAGTACAGCGCGGTGTTTGGCGCTGCGCGAGT 31

```

```

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)883-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptz9pt-Fls
; US-08-232-463-14

```

```

Query Match
Best Local Similarity 2.7%; Score 32.2; DB 1; Length 7218;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```

```

Matches 22; Conservative 205; Mismatches 186; Indels 0; Gaps 0;

QY 759 CAGCGTGCATATGAGGGTAACGTCGTCATGCTTCAATATGCAATCGGGCCAGTCG 818
Db 1435 CRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1376
QY 819 AAACCGGTCGAATACCGTTACCGGTCTCTGACACCTTCGCGCATCAGATCGCCATC 878
Db 1375 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1316
QY 879 ATGGGTGAGATAGGACAAATGAAATCAGACATATTCAGGGAACGTTCTGTCGTC 938
Db 1315 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1256
QY 939 TTACGGGTGATGTAGTGTCTGACAAATGTAAGCGTGAAGGAGTGAAGGCTGTCTCTGAAAGGA 998
Db 1255 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1196
QY 999 CTCAGTCTGTATATTCATGTCAGCCACTATGCGAGATGACGAGATCAGGTTAAT 1058
Db 1195 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1136
QY 1059 TCCCGGTAATCGTCTGAAGTCTGAGGATGGAAGGAGTGAAGGCTGTCTCTGAAAGGA 1118
Db 1135 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1076
QY 1119 ATAAAGTGACATCATCCCTCTTTTCTGGCTTCGCGAGCAATTTTACTTTTTT 1173
Db 1075 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1021

```

```

RESULT 12
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)883-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgt-Fls
US-08-232-463-14

Query Match      2.7% Score 31.8; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 6.8;
Matches 15; Conservative 204; Mismatches 176; Indels 0; Gaps 0;

QY 328 GCAGTGGCGGTCCATGACAGACATCCCGGTAACAGCGTCCCTGTACATCT 387
DB 1045 GCAGTGGCGGAGCTTGCATATTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1104

QY 388 TCTGAATGACATCAGGATCCCGCTCTACTGCGGTAACGGGACCCCGAGACTG 447
DB 1105 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1164

QY 448 ACCTTCAGCAGTACATACCAACGCTTCATTTCCGAAGGATGATCACCACACTGG 507
DB 1165 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1224

QY 508 CAATCCGCTAGACGGTAACGCTGGGAAAGGACCTGCATTAACACATCTCCGCTCA 567
DB 1225 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1284

QY 568 TTCCAGGTGTTCTGCTGCTGACGACAGCTGCTGATTTCTACGCCCGCGCCCA 627
DB 1285 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1344

QY 628 CCAGGACCGAGCAATGATTCCTTCATCTCAGCTGATACATACAGCAGCATAA 687
DB 1345 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1404

QY 688 ATTCATGCTCTTTTCGGGAGTAGCATCCCAACC 722
DB 1405 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACC 1439

```

```

RESULT 13
US-09-036-987A-1
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Query Match      2.7% Score 31.4; DB 3; Length 80161;
Best Local Similarity 56.2%; Pred. No. 37;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 563 GCTCATTTCCAGGTGTTCTGCTGCTGACGACGAGTGCTTCTTATTTTCACGCCCGGC 622
DB 62413 GCCGAGTGCCAGGTGCGCGACTGCCGGTGAATTGGGAACAGGTATTTCTCAACACCGGA 62472

QY 623 GCCCACCACGACGACGAGAAATGATTTCCCTTCCATCTTCAGCTG 667
DB 62473 GCCCGACGCGTCCGCTGCCGACCTACCCGTTCCAGCGGACGCG 62517

```

```

RESULT 14
US-09-370-700-1
; Sequence 1, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

```

```

Query Match      2.7% Score 31.4; DB 4; Length 80161;
Best Local Similarity 56.2%; Pred. No. 37;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 563 GCTCATTTCCAGGTGTTCTGCTGCTGACGACGAGTGCTTCTGTTATTTTCACGCCCGGC 622
DB 62413 GCCGAGTGCCAGGTGCGCGACTGCCGGTGAATTGGGAACAGGTATTTCTCAACACCGGA 62472

QY 623 GCCCACCACGACGACGAGAAATGATTTCCCTTCCATCTTCAGCTG 667
DB 62473 GCCCGACGCGTCCGCTGCCGACCTACCCGTTCCAGCGGACGCG 62517

```

```

RESULT 15
US-08-353-550-7
; Sequence 7, Application US/08353550
; Patent No. 5744313
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.

```



```

; LENGTH: 3031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: Shc Proteins
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..1503
; US-08-664-962B-3

Query Match          2.6%; Score 31.2; DB 4; Length 3031;
Best Local Similarity 54.3%; Pred. No. 6.5;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 358 CCACGGGTAAACAGCGTCCCTGTACACATTCTTGAATGACATCAGGATCCCGCGGTCT 417
Db 591 CCTGTGAATCAGAGAGCTGCCACATTCTGGAGTCTCCGGAAGGCTTGCACGAGTGT 650

QY 418 CACTGGCGATACAGCGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAAC 473
Db 651 CATCAGCACCATTGGCCAGGCTTCGAGTTGCGCTTCAAAACAATACCTCAGGAACC 706

RESULT 18
US-09-311-743-3
; Sequence 3, Application US/09311743
; Patent No. 6238903
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/311,743
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7771-32
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-364-7311
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: Shc Proteins
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..1503
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-311-743-3

```

```

Query Match          2.6%; Score 31.2; DB 4; Length 3031;
Best Local Similarity 54.3%; Pred. No. 6.5;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 358 CCACGGGTAAACAGCGTCCCTGTACACATTCTTGAATGACATCAGGATCCCGCGGTCT 417
Db 591 CCTGTGAATCAGAGAGCTGCCACATTCTGGAGTCTCCGGAAGGCTTGCACGAGTGT 650

QY 418 CACTGGCGATACAGCGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAAC 473
Db 651 CATCAGCACCATTGGCCAGGCTTCGAGTTGCGCTTCAAAACAATACCTCAGGAACC 706

RESULT 19
PCT-US95-03747-1/c
; Sequence 1, Application PC/TUS9503747
; GENERAL INFORMATION:
; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
; TITLE OF INVENTION: Brevican, A Glial Cell Proteoglycan
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03747
; FILING DATE: 27-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Indira, Richard J.
; REGISTRATION NUMBER: 37,643
; REFERENCE/DOCKET NUMBER: FP-LJ 1453
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 112..2848
; PCT-US95-03747-1

Query Match          2.6%; Score 31.2; DB 5; Length 3259;
Best Local Similarity 51.4%; Pred. No. 6.8;
Matches 72; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 408 CGCCCGTCTCACTGGCGATACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATATA 467
Db 2024 CTGGCACTGTCACTGGGCGAGCTGGTGGGACGCGAGCTGCTGGGACAGTT 1965

QY 468 CCAACGCTTCAATTTTCCGAAGGATGACACACACACTGCGCAATCCGGTAGACCGGTAAC 527
Db 1964 CTCTTGAATCTCTTCAGAGGGGGTCTCCAGATCCCTGTATCCCCAGGGCCGTGTG 1905

QY 528 GCTGGGAAAGGCGACCTGC 547
Db 1904 GCTGGAAGCAGGGAAGGGC 1885

RESULT 20
US-08-896-449A-1/c

```

Sequence 1, Application US/08896449A

Patent No. 6040143
GENERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: Yuzbasiyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
TITLE OF INVENTION: FACTOR AND METHODS OF USE
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS: 11

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: 5445 Corporate Drive

CITY: Troy

STATE: Michigan

COUNTRY: USA

ZIP: 48098

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/896,449A

FILING DATE: 18-JUL-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Dean F.

REFERENCE/DOCKET NUMBER: 2115-001226

TELECOMMUNICATION INFORMATION:

TELEPHONE: 248-641-1600

TELEFAX: 248-641-0270

TELEX: 287637

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 8802 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 203...8641

OTHER INFORMATION: /function= "Blood Clotting Protein"

OTHER INFORMATION: /product= "Canine von Willebrand Factor"

OTHER INFORMATION: /standard_name= "vWF"

PUBLICATION INFORMATION:

AUTHORS: Venta, Patrick J.

AUTHORS: Li, Jianping

AUTHORS: Yuzbasiyan-Gurkan, Vilma

AUTHORS: Schall, William D.

AUTHORS: Brewer, George J.

TITLE: Von Willebrand's Disease in the Scottish

TITLE: Terrier is Caused by a Single Base Deletion in

TITLE: Exon Four of the von Willebrand Factor Gene

JOURNAL: Journal of the American Veterinary Medicine Association

DATE: 1996

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 8802

US-08-896-449A-1

Query Match

Best Local Similarity 2.6%; Score 31.2; DB 3; Length 8802;

Matches 78; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 390 TGAATGACATCAGGATCCCGCCGTCCTCACTGGCGATACGGCGCAGCGGAGACTGAC 449

DB 7659 TGAAGGACATAAGTGAAGCCTTGACAGGCGAGTTGTCTCCACAGGGCTTCTGGGAGCACTGG 7600

QY 450 GCTTCAGCGCAGTACCATACCAACGCTTCATTTCCGAAGGCATGACCAACACACTGGCA 509

DB 7599 GCCACAGCGCCCATCATCAGAGTCTCCCAAGTCCGTCGAGGTGCACAGCTCACAGGCC 7540

QY 510 ATCCGGTAGACCGGTAAACGCTGGGAAGAAGGCGACCT 545

DB 7539 TCCTCCCAAGAACTGCCCCACAGGGTAGATGGTGCCT 7504

RESULT 21

US-09-132-652-1/c

Sequence 1, Application US/09132652

Patent No. 6074832

GENERAL INFORMATION:

APPLICANT: Venta, Patrick J

APPLICANT: Yuzbasiyan-Gurkan, Vilma

APPLICANT: Schall, William D

APPLICANT: Brewer, George J

APPLICANT: Duffendack, John

TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS

TITLE OF INVENTION: OF USE

FILE REFERENCE: 2115S-001226CPB

CURRENT APPLICATION NUMBER: US/09/132,652

CURRENT FILING DATE: 1998-08-11

EARLIER APPLICATION NUMBER: 08/896,449

EARLIER FILING DATE: 1997-07-18

NUMBER OF SEQ ID NOS: 29

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 1

LENGTH: 8802

TYPE: DNA

ORGANISM: Canis familiaris

US-09-132-652-1

Query Match

Best Local Similarity 2.6%; Score 31.2; DB 3; Length 8802;

Matches 78; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 390 TGAATGACATCAGGATCCCGCCGTCCTCACTGGCGATACGGCGCAGCGGAGACTGAC 449

DB 7659 TGAAGGACATAAGTGAAGCCTTGACAGGCGAGTTGTCTCCACAGGGCTTCTGGGAGCACTGG 7600

QY 450 GCTTCAGCGCAGTACCATACCAACGCTTCATTTCCGAAGGCATGACCAACACACTGGCA 509

DB 7599 GCCACAGCGCCCATCATCAGAGTCTCCCAAGTCCGTCGAGGTGCACAGCTCACAGGCC 7540

QY 510 ATCCGGTAGACCGGTAAACGCTGGGAAGAAGGCGACCT 545

DB 7539 TCCTCCCAAGAACTGCCCCACAGGGTAGATGGTGCCT 7504

RESULT 22

US-09-403-768-7/c

Sequence 7, Application US/09403768

Patent No. 6444804

GENERAL INFORMATION:

APPLICANT: Lam, Joseph S.

APPLICANT: De Kievit, Teresa R.

APPLICANT: Burrows, Lori L.

APPLICANT: Walsh, Andrew

APPLICANT: Matewish, Mauricia

TITLE OF INVENTION: No. 6444804el Proteins Involved in the Synthesis and Assembly

TITLE OF INVENTION: of Core Lipopolysaccharide of Pseudomonas Aeruginosa

FILE REFERENCE: 6580-177

CURRENT APPLICATION NUMBER: US/09/403,768

CURRENT FILING DATE: 1999-11-02

PRIOR APPLICATION NUMBER: U.S. 60/045,418

PRIOR FILING DATE: 1997-05-02

PRIOR APPLICATION NUMBER: U.S. 60/046,149

PRIOR FILING DATE: 1997-05-09

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 7

LENGTH: 1122

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa rfag (waaG)

```
;
;
; OTHER INFORMATION: Description of Organism: Pseudomonas aeruginosa
; OTHER INFORMATION: serotype O5 strain PAO1
; US-09-403-768-7

Query Match      2.6%; Score 30.8; DB 4; Length 1122;
Best Local Similarity 50.7%; Pred. No. 4.9;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 363 GGTACACAGCGTCCTGTACATTTCTTGAATGACATCAGGATCCGCCGCTCTACTG 422
Db 956 GGCAGCACCGCGCGCGCTCGCGCTAGTGGCATAGCGGACATCGGTC 897
QY 423 GCATACAGCGCGCGGAGACTGAGCTTACCGAGTACCATACCAACGCTTCTATT 482
Db 896 ACCAACACCGCGCGCGGAGAGACGCGCTCCAGCAGCAGCGTACCGGTCTCTGTTG 837
QY 483 TCCGAAGGCATGACCCACACACTGGC 508
Db 836 TAGCGCGGTGGATCAGCAGGTGGC 811

RESULT 23
US-08-961-527-155
; Sequence 155, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; OPERATING SYSTEM: HP Vectra 486/33
; SOFTWARE: MS DOS version 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 155:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-155

Query Match      2.6%; Score 30.8; DB 4; Length 6474;
Best Local Similarity 47.4%; Pred. No. 14;
Matches 92; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 351 ACATCCCCCAGGGTAACAGCGTCCTGTGCATTTCTTGAATGACATCAGGATCCCG 410
Db 5601 ACATCTTCCTCTCAATAAAATCCATCAACCCATGTTCAATAATTTTCACTTAACCCG 5660
QY 411 CCGCTCTCACTGGGATACGGGACGCGGAGACTGACCGTTCAGCCAGTACCATACCA 470
```

```
Db 5661 CCACATTAAGTAAACCGGAGTTCTTGTGACATTGACTCTAAACACACATAGGT 5720
QY 471 AAGCGTTCAATTTCCGAAGCATGACCACCACTGGCAATCCGGTAGACCGGTAACGCT 530
Db 5721 ATTCCCTCTCTATCAGAAGGAATATACATAAATCCGATATTGGTAAACTATAGTAGCT 5780
QY 531 GGGAAAGGGCACC 544
Db 5781 GGATAGATTTCACC 5794

RESULT 24
US-07-737-736B-5/c
; Sequence 5, Application US/07737736B
; Patent No. 5260199
; GENERAL INFORMATION:
; APPLICANT: DeLuca, Hector F.
; APPLICANT: Ross, Troy K.
; APPLICANT: Frahl, Jean M.
; TITLE OF INVENTION: Method Of Producing
; TITLE OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/737,736B
; APPLICATION NUMBER: 19910730
; FILING DATE: 435
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Carl R.
; REGISTRATION NUMBER: 29,437
; REFERENCE/DOCKET NUMBER: 96-296-2185-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 414-277-5715
; TELEFAX: 414-277-5774
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; AUTHORS: Baker, Andrew R.
; AUTHORS: McDonnell, Donald P.
; AUTHORS: Hughes, Mark
; AUTHORS: Crisp, Tracey M.
; AUTHORS: Mangelsdorf, David J.
; AUTHORS: Haussler, Mark R.
; AUTHORS: Pike, J. W.
; AUTHORS: Shine, John
; AUTHORS: O'Malley, Bert W.
; TITLE: Cloning and expression of full-length CDNA
; TITLE: encoding human vitamin D receptor
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 85
; PAGES: 3294-3298
; DATE: May-1988
```


RESULT 29
US-08-765-907A-1
; Sequence 1, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAMAS-JACQUES, Nathalie
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain
; APPLICANT: PARIS, Jean-Marc

Query Match	2.5%;	Score	29.6;	DB 5;	Length	4608;			
Best Local Similarity	46.6%;	Pred. No.	28;						
Matches	95;	Conservative	0;	Mismatches	109;	Indels	0;	Gaps	0;
273	TCCTCTGGCGCTGGCTGCCATCATCCGGGAAGCGTTCCGCTCGGGGATAAAAAATCCGCAGT	332							
3657	TCGTGGAGAGTGCCTGGCTGCACGCCGAGCCCGCCGCGCAAGCTCATGATTCCCATGGA	3716							
333	GGCGCGGTCCATGCGAGACACATCCCCACGGGTACACGCTCCCTGTGCACATTCTTCTGA	392							
3717	TGCGCGAGTCCCAACATCTGCTGTGCTGNGCGCCATCCGGTGCACAGCTAGAAAGTGC	3776							
393	ATGACATCAGGGATCCGGCCCGCTTCATCTGCGGATACGGGGACGCCGGAGACTGACGCT	452							
3777	CCAGTACCCAGGAATCCTCCGCTGCCACCTGTGGATATCCCCACCCGCGATTTCACCTCT	3836							
453	TCAGCCAGTACCATTACCAACGCT	476							
3837	CCGGCCTGTGCCATTCCCAACACT	3860							

```

RESULT 32
US-09-482-180A-3
; Sequence 3, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZN5SP6
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482.180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: fastseq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: variation
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc_feature
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n = A,T,C or G
; US-09-482-180A-3

```

Query Match	2.5%;	Score 29.4;	DB 4;	Length 1134;
Best Local Similarity	21.7%;	Pred. No. 14;		
Matches	66;	Conservative 53;	Mismatches 185;	Indels 0;
				Gaps 0

y	559	CTCGGCTCAATCCGAGGTGTTCTGTGCTGACGACAGAGTGTCTGCTATCTTCACGCC	618
b	200	CNCCNCCNACNCCNMGNCAYSWSNMGTGYCCNCCNAAAYCAYACNGTNSNMNSGCMNSNY	259
y	619	CGGGCCCCACGACGACGCCAGGAATGATTTCCCTTCCTATCTTCAGCTGTACAAATACAC	678

Db 260 TNSNTNCCNWSNMGNCAYMGNTNTTYTNNACNTAYMNCAYTCYMGNAAYTTWSNA 319
QY 679 GCAGCATAAATTCATGCTCTTTTCGGGAGGTAGCATCCCACTGAACGATAGCGGAA 738
Db 320 THTYNTNGARCCNWSNGGNTGYWSNAARGAYACNTTYTNTYNTNGCNATHAARWSNC 379
QY 739 CATGCTGCTGATGACGCCAGCGCTGGATATGACGGGTAAACGGTCCGATGCTTCAT 798
Db 380 ARCCNGNCAIYNGARMGNGCNGCNATHMGNWSNACNTGGGNGMNGTNGGNGNT 439
QY 799 TATGCAATGGGCGGAGTGAACCCGGTGGATAACCCGTTACCGGTGCTCTGACACCTT 858
Db 440 GGCNMGNGMGNCAIYNTAARYTNGTNTTYTNTNGGNGTNGCNGGWSNCCNC 499
QY 859 CCGC 862
Db 500 CNGC 503

RESULT 33

US-08-572-951-3/c
; Sequence 3, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: MODIFICATION OF STARCH
; TITLE OF INVENTION: SYNTHESIS IN PLANTS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: Pillsbury Madison & Sutro LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,951
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,602
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/263,921
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul N. Kokulis
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222957/1.02.15C
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2380 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-572-951-3

Query Match 2.5%; Score 29.2; DB 1; Length 2380;

Best Local Similarity 28.9%; Pred. No. 26;
Matches 110; Conservative 42; Mismatches 222; Indels 6; Gaps 1;
QY 398 ATCAGGATCCCGCCGCTCTCACTGCGGATAAGCGGCACCGGAGACTGACGCTTCAGC 457
Db 1926 RTCNCKNARNCCNCCNACNCGTCGNACNCGNACNCGTNCCTTANGCCATNGRTANAR 1867
QY 458 CAGTACATACCAAAAGCTTCATTTCCGAAGGATGACCACTGCAATGCCGTA 517
Db 1866 YTGRTTNARCCRCANGYTCRAANKNSWGGCATNARNADATRTCNCGNCCNGT 1807
QY 518 GACCGGTAAAGCTGGGAAAGGACCTGCGCATTAACACATCTCCGCTCATTTCCAGGTG 577
Db 1806 DATNCKRTGNCNARNGNACNSWRAANCCNACCCANGCNCKNACYTTRTCNSWRTGTC 1747
QY 578 TTCTGTCTGCTGACGACGAGCTGCTGTATTTTTCACGCCGCGCGCCACCGAGGCCA 637
Db 1746 NSWYTCRAANKNCKNARCATRTCTCNARRTCNCGNCKNCCNCCNARCATNACNAR 1687
QY 638 GCGAAATGATTTCCCTTCCATCTTCAGCTGATACATACAGGAGCAATAATTCATGTC 697
Db 1686 YTGNACTCTGTGNCNGCDAT-----CCARTGDTNGCRTCNGCDATDATRTCNACNCC 1633
QY 698 TTTTTCGGGAGCTAGCATCCCACTGGAACGATAAGCGGACATGTCTGCTGATGCGAGC 757
Db 1632 YTTTGRTGTTCNARCNCKNCCDTRAANCCDTRAANCGGNACRTCTCNCKNACYTGNAR 1573
QY 758 CCAGGCTGTGATATGCGAGG 777
Db 1572 NCCNARYTGNCYTGNAAG 1553

RESULT 34

US-08-832-883-48/c
; Sequence 48, Application US/08832883
; Patent No. 5807681
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; APPLICANT: Baldi, Alphonso
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS
; TITLE OF INVENTION: OF CANCER
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,883
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-13 US1
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3865 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-832-883-48

Query Match 2.5%; Score 29.2; DB 1; Length 3865;
Best Local Similarity 51.5%; Pred. No. 34;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 394 TGACATCAGGATCCCGCCGCTCTCACTGGGATAACGGGACGCCGGAGACTGACGCTT 453
DB 1176 TGAATCAGGTGATTCGGCTCGCTCAGCCTCATAAAGTCTGGGATTACAGGGTGAGC 1117

QY 454 CAGCCAGTACCATACCAACAGCTTCATTTCCGAAGGCATGACACACACACTGGCAATCC 513
DB 1116 CAACCCGCTGGCTCAGATGCTTCATTTTAAATAGTCAATCCCAAACTGTAAAGTA 1057

QY 514 GGTAGACCGG 523
DB 1056 GCTAGACCAG 1047

RESULT 35
US-08-832-877-48/c
; Sequence 48, Application US/08832877
; Patent No. 5840506
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,877
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-13 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3865 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-832-877-48

Query Match 2.5%; Score 29.2; DB 2; Length 3865;
Best Local Similarity 51.5%; Pred. No. 34; Mismatches 63; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 394 TGACATCAGGATCCCGCCGCTCTCACTGGGATAACGGGACGCCGGAGACTGACGCTT 453
DB 1176 TGAATCAGGTGATTCGGCTCGCTCAGCCTCATAAAGTCTGGGATTACAGGGTGAGC 1117

QY 454 CAGCCAGTACCATACCAACAGCTTCATTTCCGAAGGCATGACACACACACTGGCAATCC 513
DB 1116 CAACCCGCTGGCTCAGATGCTTCATTTTAAATAGTCAATCCCAAACTGTAAAGTA 1057

QY 514 GGTAGACCGG 523
DB 1056 GCTAGACCAG 1047

Db 1056 GCTAGACCAG 1047

RESULT 36
US-09-293-505-8
; Sequence 8, Application US/09293505
; Patent No. 6348575
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 8
; LENGTH: 4004
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-293-505-8

Query Match 2.5%; Score 29.2; DB 4; Length 4004;
Best Local Similarity 62.2%; Pred. No. 35;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 411 CCGCTCTCACTGGCGATAACGGGACGCCGGAGACTGACGCTTCAGCCAGTACCA 470
DB 3577 CAGTGGCTCTGGTGACGACGGGACCCCGGGGAGGACCAATCAGCTGATTCAGTATTC 3636

QY 471 AACGCTTCATTTTC 484
DB 3637 AACACATATTTGTC 3650

RESULT 37
5405943-1/c
; Patent No. 5405943
; APPLICANT: COMINGS, DAVID E.
; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED
; BEHAVIORS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/562,596
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 125,577
; FILING DATE: 25-NOV-1987
; APPLICATION NUMBER: 271,653
; FILING DATE: 16-NOV-1988
; APPLICATION NUMBER: 410,831
; FILING DATE: 22-SEP-1989
; SEQ ID NO:1
; LENGTH: 748
5405943-1

Query Match 2.5%; Score 29; DB 6; Length 748;
Best Local Similarity 51.1%; Pred. No. 15;
Matches 68; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1013 ATTACTGTCAATCGGCCACTATGCGAGATGACGAGTAAATCCCGGATAATCCG 1072
DB 345 ATCTCTGACATCGCGATTTCTGAAAGATCTCTGAAACAGATCCCAACTCCAGAGGATTCG 286

QY 1073 TCGAAGTCTGAGAGTGGAGAGGCTGTTCTCTGAAGAGTAATAAAGTGACATC 1132
DB 285 CTAAACACAGAGTTCATAGCTTGAATGATTAATAAAGATGTTTCATCATGATTTT 226

QY 1133 ATGCCCTCTTTT 1145
DB 225 ATTCTCTTTGTT 213

RESULT 38

US-08-858-207A-78/c
 ; Sequence 78, Application US/08858207A
 ; Patent No. 6348328
 ; GENERAL INFORMATION:
 ; APPLICANT: Black, Michael
 ; APPLICANT: Hodgson, John
 ; APPLICANT: Knowles, David
 ; APPLICANT: Nicholas, Richard
 ; APPLICANT: Stodola, Robert
 ; TITLE OF INVENTION: No. 6348328el Compounds
 ; NUMBER OF SEQUENCES: 552
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Smithline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/858,207A
 ; FILING DATE: 09-MAY-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/017670
 ; FILING DATE: 14-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Glimi, Edward R
 ; REGISTRATION NUMBER: 38,891
 ; REFERENCE/DOCKET NUMBER: P50475
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-4478
 ; TELEFAX: 610-270-5090
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 78:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1426 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-858-207A-78

Query Match 2.5%; Score 29; DB 4; Length 1426;
 Best Local Similarity 55.4%; Pred. No. 22;
 Matches 56; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

RESULT 39

US-08-961-527-50
 ; Sequence 50, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville

STATE: Maryland
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,527
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5338 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-50

Query Match 2.5%; Score 29; DB 4; Length 5338;
 Best Local Similarity 55.4%; Pred. No. 48;
 Matches 56; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

RESULT 40

US-09-453-702B-39/c
 ; Sequence 39, Application US/09453702B
 ; Patent No. 6365723
 ; GENERAL INFORMATION:
 ; APPLICANT: Blattner, Frederick R.
 ; Burland, Valerie
 ; Perna, Nicole T.
 ; Plunkett, Guy
 ; Welch, Rod
 ; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Quarles & Brady
 ; STREET: 1 South Pinckney Street
 ; CITY: Madison
 ; STATE: WI
 ; COUNTRY: US
 ; ZIP: 53701-2113
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 8.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/453,702B
 ; FILING DATE: 03-Dec-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/110,955

```

;
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25165
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-453-702B-39

Query Match 2.5%; Score 29; DB 4; Length 25165;
Best Local Similarity 58.8%; Pred. No. 1.2e+02;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 380 CACATCTCTTCGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGGCGCGCC 439
DB 4215 CACATTATTCAGGTCAGGTCGAGGATATCGCGCGCTGAGCGCGCAGTAACTGCGCGCT 4156
QY 440 GGAGACTCAGCTTCAGGCATGACC 464
DB 4155 GATGATTTGCGCGCTGAGCATGTTCC 4131

RESULT 41
US-08-578-634C-2
; Sequence 2, Application US/08578634C
; Patent No. 6025163
; GENERAL INFORMATION:
; APPLICANT: Vladimir Shamanin
; APPLICANT: Ethel Michele De Villiers
; APPLICANT: Zur Hausen
; APPLICANT: Irene Leigh
; APPLICANT: Harald Zur Hausen
; TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,634C
; FILING DATE: 26-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8484-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 493-4935
; TELEFAX: (650) 493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 .. 666
;
;
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 .. 666
;
US-08-578-634C-2

Query Match 2.4%; Score 28.6; DB 3; Length 668;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 910 ACAGATAATTCAGGGAARACGTTCTTACGGGTGATGTAGGTTTTTTTCTCTGACAA 969
DB 240 ACTGGTAAATACTGTAATAGAAGATGGAGATATGTTGGATATAGGTTTGGAAATATAA 299
QY 970 TAGTGAAGCGGTGACAGCATATCAGACGGCT 1000
DB 300 TAATRAAACCCCTGTCAGCAATAAGTCAGAT 330

RESULT 42
US-09-430-010-2
; Sequence 2, Application US/09430010
; Patent No. 6395512
; GENERAL INFORMATION:
; APPLICANT: Vladimir Shamanin
; APPLICANT: Ethel Michele De Villiers
; APPLICANT: Zur Hausen
; APPLICANT: Irene Leigh
; APPLICANT: Harald Zur Hausen
; TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,010
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/578,634
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8484-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 493-4935
; TELEFAX: (650) 493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 .. 666
;
;
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1 .. 666
;
US-08-578-634C-2
```


RESULT 47
US-09-575-574-3
; Sequence 3, Application US/09575574
; Patent No. 6376750
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Chao, Yu-Chan
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
; FILE REFERENCE: 08919-047001
; CURRENT APPLICATION NUMBER: US/09/575,574
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3572
; TYPE: DNA
; ORGANISM: Oryza sativa


```

US-08-279-058B-13/c
; Sequence 13, Application US/08279058B
; Patent No. 5668004
; GENERAL INFORMATION:
; APPLICANT: Michael E. O'Donnell
; TITLE OF INVENTION: DNA POLYMERASE III
; TITLE OF INVENTION: HOLOENZYME
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:

```

Db 574 GCATGGCGCAAGTAATGCATC 553

Search completed: June 19, 2003, 03:36:30
Job time : 73.8861 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 04:53:09 ; Search time 298.567 Seconds
(without alignments)
5804.490 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctccagagatgagaaaaa.....ttttactttttctgtcag 1181

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	49	4.1	1135	10	US-09-974-300-1695
c 2	48.6	4.1	783	10	US-09-974-300-6118
c 3	46.2	3.9	783	10	US-09-974-300-6111
c 4	41.6	3.5	1227	10	US-09-974-300-1771
c 5	40.6	3.4	129908	9	US-10-270-875-1
c 6	40.6	3.4	129908	9	US-10-270-878-1
c 7	40.6	3.4	129908	9	US-10-270-786-1
c 8	40.6	3.4	129908	9	US-10-270-710-1
c 9	40.6	3.4	129908	9	US-10-270-859-1
c 10	37	3.1	1108	10	US-09-974-300-1730
c 11	36.6	3.1	598	10	US-09-825-561A-7
c 12	36.2	3.1	598	10	US-09-864-761-14748
c 13	36.2	3.1	741	10	US-09-825-561A-66
c 14	36	3.0	1427	10	US-09-974-300-757
c 15	35.4	3.0	480	9	US-09-918-995-8152
c 16	35	3.0	1135	10	US-09-974-300-1702
c 17	34.8	2.9	513	9	US-10-123-155-536
c 18	34.8	2.9	802	9	US-10-184-644-312
c 19	34.8	2.9	802	9	US-10-184-634-312

20	34.2	2.9	74962	9	US-10-274-974-3	Sequence 3, Appl1
c 21	33.8	2.9	527	9	US-10-184-644-522	Sequence 522, App
c 22	33.8	2.9	527	9	US-10-184-634-522	Sequence 522, App
c 23	33.8	2.9	1686	12	US-10-044-090-362	Sequence 362, App
c 24	33.8	2.9	1712	10	US-09-880-107-3347	Sequence 3347, App
c 25	33.2	2.8	716	9	US-10-123-155-96	Sequence 96, Appl
c 26	33.2	2.8	777	9	US-10-184-644-348	Sequence 348, App
c 27	33.2	2.8	777	9	US-10-184-634-348	Sequence 348, App
c 28	33	2.8	690	9	US-10-123-155-306	Sequence 306, App
c 29	32.8	2.8	429	9	US-09-918-995-5997	Sequence 5997, App
c 30	32.6	2.8	617	9	US-10-198-846-9370	Sequence 9370, App
c 31	32.4	2.7	811	9	US-10-184-644-414	Sequence 414, App
c 32	32.4	2.7	811	9	US-10-184-634-414	Sequence 414, App
c 33	32.4	2.7	11649	9	US-09-764-891-6047	Sequence 6047, App
c 34	32.4	2.7	11649	9	US-10-091-438-272	Sequence 272, App
c 35	32.2	2.7	468	9	US-10-123-155-90	Sequence 90, Appl
c 36	32.2	2.7	1254	9	US-09-738-626-451	Sequence 451, App
c 37	32.2	2.7	1689	10	US-09-917-800A-1734	Sequence 1734, App
c 38	32.2	2.7	3309400	9	US-09-738-626-1	Sequence 1, Appl1
c 39	32	2.7	1311	10	US-09-815-242-4609	Sequence 4609, App
c 40	32	2.7	1563	10	US-09-815-242-8107	Sequence 8107, App
c 41	31.8	2.7	653	9	US-10-184-644-402	Sequence 402, App
c 42	31.8	2.7	653	9	US-10-184-634-402	Sequence 402, App
c 43	31.8	2.7	1049	9	US-10-123-155-358	Sequence 358, App
c 44	31.8	2.7	1146	9	US-09-891-641-35	Sequence 35, Appl
c 45	31.8	2.7	3153	9	US-09-759-130B-343	Sequence 343, App
c 46	31.8	2.7	3153	9	US-10-189-123-73	Sequence 73, Appl
c 47	31.6	2.7	671	9	US-10-184-644-346	Sequence 346, App
c 48	31.6	2.7	671	9	US-10-184-634-346	Sequence 346, App
c 49	31.6	2.7	2199	10	US-09-381-624A-5	Sequence 5, Appl1
c 50	31.6	2.7	2817	10	US-09-381-624A-4	Sequence 4, Appl1
c 51	31.4	2.7	776	10	US-09-902-331-1	Sequence 1, Appl1
c 52	31.2	2.6	592	9	US-10-198-846-9001	Sequence 9001, App
c 53	31.2	2.6	2176	9	US-09-843-250-8	Sequence 8, Appl1
c 54	31.2	2.6	3664	10	US-09-974-298-23	Sequence 23, Appl
c 55	31.2	2.6	2774	10	US-09-880-107-3416	Sequence 3416, App
c 56	31.2	2.6	8802	10	US-09-986-900-1	Sequence 1, Appl1
c 57	31.2	2.6	14462	9	US-09-843-250-9	Sequence 9, Appl1
c 58	31.2	2.6	32192	9	US-10-092-154-1416	Sequence 1416, App
c 59	31.2	2.6	32192	9	US-09-764-891-7945	Sequence 7945, App
c 60	31.2	2.6	32192	9	US-09-764-891-8220	Sequence 8220, App
c 61	31.2	2.6	32192	9	US-09-764-891-8319	Sequence 8319, App
c 62	31.2	2.6	32192	10	US-09-764-847-1416	Sequence 1416, App
c 63	31.2	2.6	34337	10	US-09-741-149-3	Sequence 3, Appl1
c 64	31	2.6	734	9	US-10-184-644-458	Sequence 458, App
c 65	31	2.6	734	9	US-10-184-634-458	Sequence 458, App
c 66	31	2.6	1440	9	US-09-891-641-47	Sequence 47, Appl
c 67	31	2.6	2652	9	US-10-195-970-4	Sequence 4, Appl1
c 68	31	2.6	2652	9	US-10-195-970-5	Sequence 5, Appl1
c 69	31	2.6	22299	9	US-10-091-572-469	Sequence 469, App
c 70	31	2.6	22299	9	US-09-764-891-6605	Sequence 6605, App
c 71	30.8	2.6	360	10	US-09-764-877-587	Sequence 587, App
c 72	30.8	2.6	491	9	US-09-918-995-15312	Sequence 15312, A
c 73	30.8	2.6	4385	9	US-10-050-704-77	Sequence 77, Appl
c 74	30.8	2.6	4385	9	US-10-050-704-78	Sequence 23, Appl
c 75	30.8	2.6	4386	9	US-10-050-704-78	Sequence 78, Appl
c 76	30.8	2.6	5511	9	US-09-974-298-13	Sequence 13, Appl
c 77	30.8	2.6	5511	12	US-10-044-090-285	Sequence 285, App
c 78	30.6	2.6	300	9	US-10-184-644-492	Sequence 492, App
c 79	30.6	2.6	300	9	US-10-184-634-492	Sequence 492, App
c 80	30.6	2.6	451	9	US-10-123-155-126	Sequence 126, App
c 81	30.6	2.6	468	9	US-10-123-155-90	Sequence 90, Appl
c 82	30.6	2.6	651	10	US-09-811-284-26	Sequence 26, Appl
c 83	30.6	2.6	705	10	US-09-974-300-2795	Sequence 2795, App
c 84	30.6	2.6	1114	10	US-09-764-864-506	Sequence 506, App
c 85	30.6	2.6	1548	9	US-10-037-270-965	Sequence 965, App
c 86	30.6	2.6	1802	9	US-09-892-877-70	Sequence 70, Appl
c 87	30.6	2.6	1802	9	US-09-948-783-70	Sequence 70, Appl
c 88	30.6	2.6	5957	10	US-10-125-540-627	Sequence 627, App
c 89	30.6	2.6	5957	10	US-09-764-870-627	Sequence 627, App
c 90	30.4	2.6	175	9	US-10-025-380-873	Sequence 873, App
c 91	30.4	2.6	175	10	US-09-922-217-873	Sequence 873, App
c 92	30.4	2.6	175	10	US-09-833-263-873	Sequence 873, App

Db 631 TTCCCAAGGGGCATATAACAGTTACTTCTGCTTCAACAACTTCCGGGATGCCACCA 572
QY 414 GTCTACTGGCGATACAGGCGACGCGGAGACTGACGCTTACGCCAGTACCATACCAAC 473
Db 571 ATTTTGTCCGATCAGAGTACCCCGCATCCGCTTCAAGCGCCACTAAACCAAG 512
QY 474 GCTTCAATTTCCGAAGGCATGACCAACACACTGGCAATCCGGTAGACCGGTACGCTGG 533
Db 511 CTTTCCTTTTCGTTTAAACACAGCAATTAATCGCTCATAGACAAATTCAGCAACATGC 452
QY 534 AAAAGGGACCTGCCATCACTCCGCTCATTCCTCCAGGCTGTTCTGCTGCTGACGC 593
Db 451 TTTTGTGCTGCTAGCATAAGGACATGTTTGTCAAGCCGTTTCTTCCACAACTCCCTC 392
QY 594 AGACGTG 600
Db 391 GCAATTG 385

RESULT 3

US-09-974-300-6111/c
; Sequence 6111, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2001-10-05
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6111
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-6111

Query Match 3.9%; Score 46.2; DB 10; Length 783;
Best Local Similarity 59.5%; Pred. No. 4.5e-05;
Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 354 TCCCCACGGGTAAACAGCGTCCCTGTACATTTCTTGAATGACATCAGGAGTCCGCCCC 413
Db 633 TTCCCAAGGGGCATATAACAGTTACTTCTGCTTCAACAACTTCCGGGATGCCACCA 574
QY 414 GTCTACTGGCGATACAGGCGACGCGGAGACTGACGCTTACGCCAGTACCATACCAAC 473
Db 573 ATTTTGTCCGATCAGAGTACCCCGCATCCGCTTCAAGCGCCACTAAACCAAG 514
QY 474 GCTTCAATTTTC 484
Db 513 CTTTCCTTTTC 503

RESULT 4

US-09-974-300-1771/c
; Sequence 1771, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression

FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2001-10-05
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1771
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1771

Query Match 3.5%; Score 41.6; DB 10; Length 1227;
Best Local Similarity 62.5%; Pred. No. 0.0024;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 352 CATCCCCACGGGTAAACAGCGTCCCTGTACATTTCTTGAATGACATCAGGATCCCGC 411
Db 1072 CATCTCTTTTAGCGGTACCATACCGGTACATTTATGTTCAACGATTTTCGGGAAGTCCCC 1013
QY 412 CGTCTCACTGGCGATAACGGGCACCGCGGAGACTGACGCTTCA 455
Db 1012 CTGCATCGCTGACGATAATAGGAAGTCCGCGCTAATTGCGCTTCA 969

RESULT 5

US-10-270-875-1/c
; Sequence 1, Application US/10270875
; Publication No. US20030082741A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arthor Aevardson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,875
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-875-1

Query Match 3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGGATAACGGGCACCGCGGAGACTGACGCTTACGC 457
Db 10181 ATCTGCAATACCACCGCTCATAGTACCAACCGGGGTCCACACAGCGACGCTCAAG 10122
QY 458 CAGTACCATCAACAAAGCTTCATTTCCGAAGGACGACCACTGACCACTGCGC 508
Db 10121 CGAAGCGATCCCGAAGCCCTCATATAGTGAAGCTGCAGCACCACTGCGC 10071

RESULT 6

US-10-270-878-1/c
; Sequence 1, Application US/10270878
; Publication No. US20030082790A1
; GENERAL INFORMATION:

```

; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-878-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071

RESULT 7
US-10-270-786-1/c
; Sequence 1, Application US/10270786
; Publication No. US20030087392A1
; GENERAL INFORMATION:
; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,786
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-786-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071

RESULT 8
US-10-270-710-1/c
; Sequence 1, Application US/10270710
; Publication No. US20030092128A1
; GENERAL INFORMATION:
; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,710
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-710-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071

RESULT 9
US-10-270-859-1/c
; Sequence 1, Application US/10270859
; Publication No. US20030092134A1
; GENERAL INFORMATION:
; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,859
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-859-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071
```

```

; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-878-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071

RESULT 7
US-10-270-786-1/c
; Sequence 1, Application US/10270786
; Publication No. US20030087392A1
; GENERAL INFORMATION:
; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,786
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-786-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071

RESULT 8
US-10-270-710-1/c
; Sequence 1, Application US/10270710
; Publication No. US20030092128A1
; GENERAL INFORMATION:
; APPLICANT: Sigrídur Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,710
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-710-1

Query Match      3.4%; Score 40.6; DB 9; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.11;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGGCGGAGACTGACGCTTCAGC 457
DB 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACACCGGGGCTCCACAGCGGCGCTCAAG 10122

QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGCGATGACACCATGCGC 508
DB 10121 CGAAGCGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCATATCGGC 10071
```

```

Db 10181 ATCTGCAATACCAACCGTCTAGTACCAACCAACCGGGCTCCACACAGCGAGCGCTCAAG 10122
Qy 458 CAGTACATACCAACCGTCTATTTCCGAAGGATGACACACACTGGC 508
Db 10121 CGAAGCATCCGAACCGCTCATAGTAGAAGAGTGCAGCACCACATCGGC 10071

```

RESULT 10

```

US-09-974-300-1730/c
; Sequence 1730, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1730
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1730

```

```

Query Match 3.1%; Score 37; DB 10; Length 1108;
Best Local Similarity 54.9%; Pred. No. 0.09;
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 387 TTCTGAATGATACATGAGGATCCCGCGTCTCACTGCGGATACAGCGGCACCGGAGACT 446
Db 926 TTCTGTATTAATCTTTCAGCGCGCGGAGTCTGCCCGATACAGCGGCTCCCGCAGCG 867
Qy 447 GAGCTTCACGCCAGTACCATACCAACGCTTCAATTTCCGAAGGATGACCCACACTG 506
Db 866 AGCGCTTCAAGGCGGGAATTTCCGAGGTTTCGTCGGAAGGAAGAAAGCAACAATCT 807
Qy 507 GCAATCCGGTAGA 519
Db 806 GAGTCCCATAGA 794

```

RESULT 11

```

US-09-825-561A-7
; Sequence 7, Application US/09825561A
; Patent No. US2002013767A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: West, James W.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Holly, Richard D.
; APPLICANT: Nelson, Andrew J.
; TITLE OF INVENTION: SOLUBLE ZALPHAL1 CYTOKINE RECEPTORS
; FILE REFERENCE: 00-22
; CURRENT APPLICATION NUMBER: US/09/825,561A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/194,731
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/222,121
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of soluble
; OTHER INFORMATION: zalphal1 Receptor polypeptide as shown in SEQ ID
; NAME/KEY: misc_feature
; LOCATION: (1)..(654)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-561A-7

Query Match 3.1%; Score 36.6; DB 10; Length 654;
Best Local Similarity 30.7%; Pred. No. 0.089;
Matches 74; Conservative 35; Mismatches 132; Indels 0; Gaps 0;

Qy 883 GTGAGATAGGCACAAATCAATCAACACAGATTAATTCAGGAAAACGTTCTGCTTTAC 942
Db 60 GTGGAAYTTCAYCCNWSNACNTNACNTGCGCARGAYCARTAYGARGARYTNA 119
Qy 943 GGGTGATAGTGGTTTCTGACAAATAGTGAAGCGGTGCACAGCATATCAGACGGCTCA 1002
Db 120 RGATGARGCNACNWSNTGYWSNTNCAIYNGNWSNCNCAAYAGCNACNACNTA 179
Qy 1003 GTCTGTCTATATTACTGTCTGTCATGCGCCTATGCGCAGATGACGAGTAAATTTCCC 1062
Db 180 YACNTGYCATGAYGTNTTYCATYTTATGCGNGAYGAYATHTTYSNGTNAAYATAC 239
Qy 1063 CGATAATCCGTCGAAGTCTGAGGATGAAGGATGAAGGCTGTTCTCTGAAAGGAATA 1122
Db 240 NGAYCARWSNGNAAATAYWSNCARGARTGYGNWSNTTYTNTNGCNGARWSNATHAA 299
Qy 1123 A 1123
Db 300 R 300

RESULT 12
US-09-864-761-14748/c
; Sequence 14748, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662

```

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 14748
LENGTH: 598
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC018525.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.48
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
US-09-864-761-14748

Query Match 3.1%; Score 36.2; DB 10; Length 598;
Best Local Similarity 56.2%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;
Matches 68; Conservative 0;

QY 1053 TTAATTCCTCCGATATCGTGAAGCTGAGAGTGAAGAGGAGTGAAGGCTGTTCCTG 1112
DB 503 TAAATGTACAGGCCACTGTGACAGAGCTCATGACAGTGAATGAACATCTATTCTC 444
QY 1113 AAAGGAATAAAGTGCATCATGCGCTCTTTTCTGGCTTCGGAGCAATTTACTTTT 1172
DB 443 AACAGATAAAGTGCATCATGACATCTCTGACAGGTTCTTGAGTTGGATATTTT 384
QY 1173 T 1173
DB 383 T 383

RESULT 13
US-09-825-561A-66
Sequence 66, Application US/09825561A
Patent No. US20020137677A1

GENERAL INFORMATION:
APPLICANT: Sprecher, Cindy A.
APPLICANT: No. US20020137677A1ak, Julia E.
APPLICANT: West, James W.
APPLICANT: Presnell, Scott R.
APPLICANT: Holly, Richard D.
APPLICANT: Nelson, Andrew J.
TITLE OF INVENTION: SOLUBLE ZALPHA11 CYTOKINE RECEPTORS
FILE REFERENCE: 00-22
CURRENT APPLICATION NUMBER: US/09/825,561A
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/194,731
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/222,121
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 86
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 66
LENGTH: 741
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Degenerate polynucleotide sequence of SEQ ID NO:69
NAME/KEY: misc_feature
LOCATION: (1)...(741)
OTHER INFORMATION: n = A,T,C or G

US-09-825-561A-66

Query Match 3.1%; Score 36.2; DB 10; Length 741;
Best Local Similarity 27.6%; Pred. No. 0.13; Mismatches 48; Indels 0; Gaps 0;
Matches 83; Conservative 48;

QY 833 AACCGTTACGGTGCTCTGACACTTCCGCCATFACATGCGCCATCATCGGTGAGATAGG 892
DB 139 AAYTTCAYCCNNSNACNTYNACNTGATHYTNWSNAAAYAYACNGGNTGYTAY 198
QY 893 CACAACAATGAATCACACAGATAAATTCAGGGAACGGTCTCTGCTTACGGGTGATGTA 952
DB 199 ATHAARGAYMGNCNTYNGAYTNMGNCARGAYCARTAYGARGARTYNAARGAYGARGCN 258
QY 953 GGTTTTGTCTGACAACTAGTGAAGCGGTGACACATATCAGACGGCTCAGTCTCTCTAT 1012
DB 259 ACNWSNTGYWSNTNCAVMGNSNCNCAAYAGCNACNCAVCNACNTAYACNTGYCAY 318
QY 1013 ATTACTGTATGCCACTATGGCAGATGACCATCATAGGTTTAAATCCCGATATCCG 1072
DB 319 ATGGAYGTNTTYCAATTTATGCGNGAYGAYATHTTYSNGNTNAYATACNGAYCARWSN 378
QY 1073 TCGAAGTCTGAGGATGAAGGAAGGTGAAGGCTGTCTCGAAAGGAATAAAAGTGACATC 1132
DB 379 GGNAAATYWSNCARGATGGNWSNTTNTTNCNCARGWSNMGNCARTAYAAVATH 438
QY 1133 A 1133
DB 439 W 439

RESULT 14

US-09-974-300-757/c
Sequence 757, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
TITLE OF INVENTION: Expression
FILE REFERENCE: 10085.500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 757
LENGTH: 1427
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-757

Query Match 3.0%; Score 36; DB 10; Length 1427;
Best Local Similarity 46.0%; Pred. No. 0.24; Mismatches 160; Conservative 0; Indels 185; Gaps 1;
Matches 160;

QY 398 ATCAGGATCCCGCCGCTCTCTCTGCGCATACGGGACGCGGAGACTGACGCTCAGC 457
DB 1215 ATCATGAAGCCGCGCTCTCTCTGCAACCGGACCGCGCGCTACCTGAGGGCAATCAG 1156
QY 458 CAGTACCATACCAACCGTTCTATTTTCGAAGGCATGACACCATGCGCAATCCGGTA 517
DB 1155 CTGTCCAGACCGCACGGCTCGAATTTTGAAGGCATCAGGAAGATCGCTCGCGGCATA 1096
QY 518 GACCGGTAACGCTGGGAAAAGGCACTGCCATTAACACATCTCCGCTCATTTCCAGGTG 577
DB 1095 AATCTTGGCGTGAAGGTTTCATGAAAACCGATATACGCTTGACCTGTCCGAAAAAGC 1036
QY 578 TTCTGTCTGCTGACGACGACGCTGCTTCGTATTTTCACGCCCGCGCCACACACGACCA 637
DB 1035 ATGCTCCATATAGCGGAAGTAATTTTCAATTTCCGCTTCTCCGGTCTCTAAACGATCAG 976

Db 279 N.BG..BS:CS.T...NCS.H.Y..SSWASKA.C.BS.MNW...R.N..SB.Y...T.N. 220
QY 543 CTOCCATTACACATCCCGCTCAATVCCAGGTGTTCTGCTGACGACGACGTGT 602
Db 219 MS..ND..D.M.Y.KCT.BNYT.SY.YNRC.H...NYGH...YB..AY.SNSY.D.S. 160
QY 603 TCGTATCTTACGCGCGGCCGCCACACGAGCCAGGGAATGATTTCCCTCCATCTTC 662
Db 159 .MY.CY...SC.SD..NRS.H.NT.NAB.Y.ANV..RS.YNSS...M.YY.CN.TN.H. 100
QY 663 AGCTGATACATACACGACGACATAAATTCATGCTCTTTTCGGGACGTAGCATCCCAACC 722
Db 99 NS.DNDH.R.WA..N.CM..NRM.M..SNRY.S.C.GCTS.SS.....M.S.GRBKMC.G 40
QY 723 TGAAGGATAACGGGA 737
Db 39 YGCM.GCY.TSS.KA 25

RESULT 18

US-10-184-644-312/c

; Sequence 312, Application US/10184644

; Publication No. US20030044930A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P34301C227

; CURRENT APPLICATION NUMBER: US/10/184,644

; CURRENT FILING DATE: 2002-06-28

; Prior Application removed - See File Wrapper or Palm

; SEQ ID NO 312

; LENGTH: 802

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-184-644-312

Query Match 2.9%; Score 34.8; DB 9; Length 802;

Best Local Similarity 5.2%; Pred. No. 0.43;

Matches 28; Conservative 170; Mismatches 336; Indels 0; Gaps 0;

QY 190 TGTGTCATTTCCCTTAGCTGACCTAGCTAGCCAGACACAAATGATCTGCGGTTCTGTAA 249
Db 591 YRHYK..B.MTMH.NSHSN...BB.MBN.N...GAM..NCK.YB.KH.YCS.....N.MM 532
QY 250 TATCAAACCGGTACTCAATATCTCTGCGCTGCGTGCATCATCCGGAAGCGTTCCG 309
Db 531 HM..MD.SMSC.....B.MM.SNS..S.SCS..N.YMT...TH...D.B...MS.A...R 472
QY 310 GTCGGGATAAAATCCGCGAGTCCGCGGTCTGATGACATCAGGATCCCGCGGTCTCACTGGCGATA 369
Db 471 TDYTB.BAKMCMCDMMR....H..R.TCRK.NBYRCR...M.H..YCYAMTNTN.K.KB 412
QY 370 GCGTCCCTGTACATTTCTTGAATGACATCAGGATCCCGCGGTCTCACTGGCGATA 429
Db 411 .KB.HH..T.Y.TSHKYY...CHTSR.RMCA...W.....Y.Y...YK.S...C...H. 352
QY 430 CGGCGACGCGGAGACTGACGTTTACGAGTACCATACCAACGCTTCACTTCCGAAG 489
Db 351 MD..ANSRMSK.B..YD...AYDY..S.MYS.KRSD.YR.R..NNN.D.A.T.D.MSND. 292
QY 490 GCATGACACACACTGGCAATCCGCTAGACCGGTAACTGGGAAAAGGACCTGCCA 549
Db 291 HY.RCMWN..RN.R...YK.RSWBG.BCT.YYBGC.BH.DAARK.Y.GMC.D.BKYY. 232

Db 291 HY.RCMWN..RN.R...YK.RSWBG.BCT.YYBGC.BH.DAARK.Y.GMC.D.BKYY. 232
QY 550 TTAACACATCTCCGCTCATTCGCCAGTGTCTGCTGCTGACGAGAGCGTCTGCTATT 609
Db 231 B.YSK.BC.CCKG.N.C...T.MCK..AAC.CA.C...SSN.SY..N..YHCM.RBHH 172
QY 610 CTTACGCGCGCGGCCGCCACACGAGCCAGGAAATGATTCCTTCCATCTTCACTGAT 669
Db 171 HHYK.W.RMH.RDHKRM.KK.SMM...RSHEBCY..BB...BSAHSK.SS....BMC..A 112
QY 670 ACAATACAGCAGCAATAATTCATGCTCTTTTCGGGACGTAGCATCCCACT 723
Db 111 MSWAYTTBTYA..RM.TAKCB...NVHY.CCH.HSC..W..TCYTHCYTCT 58

RESULT 19

US-10-184-634-312/c

; Sequence 312, Application US/10184634

; Publication No. US20030068684A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P34301C217

; CURRENT APPLICATION NUMBER: US/10/184,634

; CURRENT FILING DATE: 2002-06-28

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 312

; LENGTH: 802

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-184-634-312

Query Match 2.9%; Score 34.8; DB 9; Length 802;

Best Local Similarity 5.2%; Pred. No. 0.43;

Matches 28; Conservative 170; Mismatches 336; Indels 0; Gaps 0;

QY 190 TGTGTCATTTCCCTTAGCTGACCTAGCTAGCCAGACACAAATGATCTGCGGTTCTGTAA 249
Db 591 YRHYK..B.MTMH.NSHSN...BB.MBN.N...GAM..NCK.YB.KH.YCS.....N.MM 532
QY 250 TATCAAACCGGTACTCAATATCTCTGCGCTGCGTGCATCATCCGGAAGCGTTCCG 309
Db 531 HM..MD.SMSC.....B.MM.SNS..S.SCS..N.YMT...TH...D.B...MS.A...R 472
QY 310 GTCGGGATAAAATCCGCGAGTCCGCGGTCTGATGACATCAGGATCCCGCGGTCTCACTGGCGATA 369
Db 471 TDYTB.BAKMCMCDMMR....H..R.TCRK.NBYRCR...M.H..YCYAMTNTN.K.KB 412
QY 370 GCGTCCCTGTACATTTCTTGAATGACATCAGGATCCCGCGGTCTCACTGGCGATA 429
Db 411 .KB.HH..T.Y.TSHKYY...CHTSR.RMCA...W.....Y.Y...YK.S...C...H. 352
QY 430 CGGCGACGCGGAGACTGACGTTTACGAGTACCATACCAACGCTTCACTTCCGAAG 489
Db 351 MD..ANSRMSK.B..YD...AYDY..S.MYS.KRSD.YR.R..NNN.D.A.T.D.MSND. 292
QY 490 GCATGACACACACTGGCAATCCGCTAGACCGGTAACTGGGAAAAGGACCTGCCA 549
Db 291 HY.RCMWN..RN.R...YK.RSWBG.BCT.YYBGC.BH.DAARK.Y.GMC.D.BKYY. 232


```

; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 96
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-96

Query Match      2.8%   Score 33.2; DB 9; Length 716;
Best Local Similarity 8.0%; Pred.No.1.4;
Matches 40; Conservative 135; Mismatches 322; Indels 0; Gaps 0;

Qy      343 ATGCAGACATCCCCACGGGTAAACAGCGTCCCTGTCAATTTCTTCTGAATGACATCAG 402
Db      :| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      541 RTA.AMNMNH..BSYV..CK..MHMS....MKAER.MCHBTM..TT.NB.N...TCS.CB 482
Db      :| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      403 GGATCCGCGCGTCTACTGCGGATACAGGCGACGCGGAGACTGACGCTTCAGCCAGTA 462
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      481 ACYT....HBS.H.M.MN.RWM...AYCMBH.Y.BAB.KH.YC.Y..T.NH.T...NATC. 422
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      463 CCATACCAACGCTTCATTTTCCGAAGCATGACCACCACTCGGCAATCCGGTAGACCG 522
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      421 ..BC.N.M.CHK.T...N.A.YSR.DK.S..YMCBSH..H...BG.TNTMT.Y..NY.Y 362
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      523 GTAACGCTGGGAAAGGCGACCTGCCATTAACACATCTCGCTCATCCAGAGGTGTTCTG 582
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      361 STCECB.K.H..SCSTRR..BHT..CTBTYT..AMCAC..CB...CM..MCC..A.M..NM 302
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      583 TCTGCTGACGCGACAGCGTCTCGTATTCTTCACCGCGGGCGCCACACGAGCCAGCGAA 642
Db      301 ...BB.....MT..BCHMD..ABNKK.B.HBTSH.CAAAY.YBS...NM..CR.C....B 242
Qy      643 ATGATTTCCTTCATCTTCAGCTGATACATAACAGCAGCATAAATTCATGCTCTTTT 702
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      241 ...YAYV.SHNAM.A.T.M..C.T.CAM.TH.DT...NMAMK.MH.S.BHSCVHY... 182
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      703 CGGACGTAGCATCCCCACCTGAACGATAGCGGAACATTCCTGCTGATGCACCCAGG 762
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      181 CMB..S.T.N.A.T.Y.SAS...YYAYSM.AM.CYS...B.WR..HSG.S..TYSDDYR.GS 122
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      763 CGTGGATATGACGGGGTAACGGTCCGCTGCGCTTCATTATGCAATGCGGGCGAGCGAAAC 822
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      121 SY.AS..H.NCRMN....TSBDSADWDSS.NWCMG.C...N.B..H..C.M..S.C.HY 62
Db      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      823 CCGGTGGAATAACCGTT 839
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      61 .N.S..SS.T.D..BBH 45
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 26
US-10-184-644-348/c
; Sequence 348, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P343OR1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-348

```

Query Match	2.8%;	Score 33.2;	DB 9;	Length 777;	
Best Local Similarity	7.5%;	Pred. No. 1.5;			
Matches	42;	Conservative 179;	Mismatches 342;	Indels 0;	Gaps 0;
Qy	177	ATCCGCGGTATATCTGTTGCTATTTCCCTTAGCTAGTACCGACAGACACAATGATCGT	236		
Db	765	DDYNYMMK..KDW..CCMNY.YYM.YDWK..GR.H.S.N.SS....RMRY.YS.T. 706			
Qy	237	GCCTTCTGTTAATATCAAAACCGGTACTACTATATCTTCTCGGCGCTGGCTGCCATCATC	296		
Db	705	.HMB.C.D.TY.AN.K.N..BN.A.MB.AD..AD..TMGRRKSHMM..SY...CR.A. 646			
Qy	297	CGGAAGCGTCCGCTCGGGATAAAAATCGCGAGTCGCGCGTCCATGACAGACACATCC	356		
Db	645	M..Y.H.M...YD.HCSY..RWM.AT..SM..G..ASN...C..BM.HTA.DS..SH..H 586			
Qy	357	CCCACGGTAACACGCTCCCTGTGCATCTTCTGAATGACATCAGGGATCCCGCCCGTC	416		
Db	585	WG.A..HCRMBH.YTTYMSA.TRYSGTNCHWTGR.HYT.CGHGTGTCRAIGYD..S..B. 526			
Qy	417	TCAGTGGGATACGGGACGCCCGGAGACTGACGCTTACGCCAGTACCATCAACAAACGCT	476		
Db	525	CHYSC.R...M.S..KN...SSDM.....BB..KNWM.MS..SBBM.BAC.HAC..KB 466			
Qy	477	TCATTTCCGAAGCATGACCACACACTGCGATCCGGTAGACCGGTACCGCTGGGAA	536		
Db	465	HR.CH.T.BDHHB..A.YRHBN.YM.A.CCTTB.RBSMKRBSDM..S.BHH..HYASM.. 406			
Qy	537	AGGCACCTGCGCATTAACACATCTCCGCTCATATCCACAGTGTTCGTGCTGACGCAGA	596		
Db	405	.HRAMS.GAC.Y.R..YCHR.BWYDHTS.MDTR.CN.BTY.HTKSRBGTSCM..SSAA 346			
Qy	597	CGTCTTCGTATTTCTACGCGCGCGGCCACAGAGCCACAGCGAAATGATTCCTTC	656		
Db	345	.BCRBB.NY.HYA..R.H..H.RAHTCHSC..SG..YTM..AAWMN..SY.CCBHNMG 286			
Qy	657	ATCTTCAGCTGATACATACAGCAGATAAATCATCTGCTCTTTTCGGGAGGTAGCATC	716		
Db	285	BYCYS..AMHASCS..SS.Y..R.MHHH.NRAH....AC..MTCN.WRD..S.HAY..RD 226			
Qy	717	CCCACCTGAAGATAAGCGGAAC	739		
Db	225	DHDA.C.SYA.TAHMC..HSTAC 203			

RESULT 27
US-10-184-634-348/c
; Sequence 348, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pap, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhen

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 346
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-348

```

Query Match	2.8%;	Score 33.2;	DB 9;	Length 777;
Best Local Similarity	7.5%;	Pred. No. 1.5;		
Matches 42;	Conservative 179;	Mismatches 342;	Indels 0;	Gaps 0;
QY	177	ATCCGCCGTTATCTGTTGGCAATTTCCCGCTTAGCCTGACTACCGCAGAGACACAATGATCTGT	236	
Db	765	DDYYNYMMK..KDWMM.CCMNY.YYM.YDWK..GR.H.S.N.SS....RHMY.YS.T.706		
QY	237	GCGGTCCTGTAATCAACACCGGTACTCAATATCTTCTGCTGCCTGCCTGCCATCATC	296	
Db	705	.HMB.C.D.TY.AN.K.N..BN.A.MB.AD..AD..TWGRKKSHMM..SY...CR.A.645		
QY	297	CGGAAGCGTTCCGTCGGGATAAAAAATCGCGCAGTGCGCCGCTCCATGACGACACATCC	356	
Db	645	M.Y.H.M...YD.HCSY..RWM.AT..SM..G...ASN...C..BM.HTA.DS.SH..H.586		
QY	357	CCACGGGTAAACAGCGTCCCTGTCCACATTTCTTGAATGACATCAGGAGTATCCGCCGCTG	416	
Db	585	WG.A..HCRMBH.YTYIYMSA.TRYSGTNCHWTGR.HYT.GGHTGTMCRAHGYD.S..B.526		
QY	417	TCATGCGGATAACGGGACCGCGAGACGTGACGCTTCAGCCAGTACCCAGTACCATCAACAACCT	476	
Db	525	CHYC.R...M.S...K.N....SSDM.....BB..KNWM.MS.SBBM.BAC.HAC..KB.466		
QY	477	TCATTTTCCGAAGGCGATGACCACACACACTCGCAATCGGTAGACCGGTAAACGCTGGGAAA	536	
Db	465	HR.CH.T.BDHBB..A.YRHBN.YM.A.CCTB.RBSMRKBSYDM..S.BHH..HYASM...406		
QY	537	AGGGACCTGCCATTACACATCTCCGCTCATTTCCAGGTGTCTGTCTGCTGACGCAGA	596	
Db	405	.HRAMS.GAC.Y.R..YCHR..BWDHYTS.WDTR.CN.BTY.HFKSRBGTSCHM...SSAA.346		
QY	597	CGTGCTTCGTATCTTCTACGCCGCCGCCGCCACACGAGCAGCGAATAATGATTCCTTCC	656	
Db	345	.BCRBB.NY.HVA...R.H...H.RAHTCHSC..SG..YTM..AAMWN..SY.CCBHNG.286		
QY	657	ANCTTCAGCTGATCAATAACACGAGCATCAATAATTCATGCTCTTTTCGGGACGTAGCATC	716	
Db	285	BYCBYS..AMHSASC..SS.Y...R.MHHH.NRAH.....AC..MTCN.WRD.S.HAY..RD.226		
QY	717	CCCACTGCATCAAGCAATAGCGGAAC	739	
Db	225	DHDA.C.SYA.TAHMC..HSTAC.203		

RESULT 28
US-10-123-155-306/c
: Sequence 306, Application US/10123155
: Publication No. US2003006879A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.

```

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 306
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-123-155-306

Query Match          2.8%; Score 33; DB 9; Length 690;
Best Local Similarity 9.8%; Pred. No. 1.7;
Matches 32; Conservative 100; Mismatches 195; Indels 0; Gaps 0

QY      314  GGATAAAAAATCGCGCAGTGGCGGGTCCCATGCGACACATCCCCACGGGTAAACAGCGT 373
      1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-: 1-:
DB      401  GSADA..NSRA..GC.SWSCKNAH.SRNW.TG.SYRVHBMVYD.S.A.M...R.A..NS 342

QY      374  CCTCTGCATCTTCTGTAATGACATCAGGATGCCCGCGTCTCACTGGCGATAACGGG 433
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      341  SKFSB..BSB.BY....SNHRN..M...NHSSS...C.SMRP...TBTBNCNSHRTTMYM... 282

QY      434  CACGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTCCGAAGGCAT 493
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      281  .N.RICHKNKD.D.DMKNRSH...BM.T.HASNAH...A.M..S.S.Y.AT..BADK.MA. 222

QY      494  GACCACCACTGGCAATCCGGTAGACGGTAAAGCTGGGAAAGGGCACCCTGCCAATAA 553
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      221  DAYYDNBS.MHWBB.AHY.B..NNBAMB..A.ASNS.AHMTS.ANNMRC..SSS.T....R 162

QY      554  CACATCTCGGTCATCTCCAGGTCTGTCTGCTGACGACGAGCTGCTCGTATTTCTTC 613
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      161  A..HA.S.HABSNYRB....TB..M.SY.MA.AMN.NTT.GBNI.DGNTNBN..GBACHN 102

QY      614  AGCGCGGCGGCCACCAACGACGCCAGCG 640
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      101  A..YH.NSSSY..C.BGKRRRC.ANAG 75

```

```

RESULT 29
US-09-918-995-5997
; Sequence 5997, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE OF INVENTION: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5997
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-5997

```

Query Match 2.8%; Score 32.8; DB 9; Length 429;
Best Local Similarity 55.2%; Pred. NO. 1.4;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 358 CCACGGGTACAGCGTCCCTGTCACATCTTCTGAATGACATCAGGGATCCCGCCGCTT 417
DB 202 CCTGTGTAATCAGAGAGCGCTGCCACATCTTGAGTGTCCCGAAGGCTTGCACGAGTGT 261
QY 418 CACTGCGGATACAGCGCAGCGGAGCAGTGCAGCTTCAGCGTAGTACCATCAACAAAC 473
DB 262 CATCAGCACCATGGCCAGCGCTTCGAGTTGCGCTTCAACAACATACCTCAGCAACC 317

RESULT 30

US-10-198-846-9370
; Sequence 9370, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9370
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15, 17, 27, 62, 165, 441, 544, 600
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9370

Query Match 2.8%; Score 32.6; DB 9; Length 617;
Best Local Similarity 54.2%; Pred. No. 2.1;
Matches 65; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 4 CAGGAGATGAAAAAGCCAAATAAATAATGCCATCCAGCGCGCTCCAGTGAAA 63
DB 370 CATGAGATTTAAAAACAAAAAAGCGCGCCACCGCGTGGAGCTCCAA 429
QY 64 GTAGGCGTCTGTGTCGCGTATTAAATGCATTGACCGTCCCGCTATTAAACAATGTA 123
DB 430 TTCGCCCTATANTGAGTCTGATTACGCGGCTCAGTGGCGCTGTTTACACGCTGTGA 489

RESULT 31

US-10-184-644-414
; Sequence 414, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 414
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-414

Query Match 2.7%; Score 32.4; DB 9; Length 811;
Best Local Similarity 7.3%; Pred. No. 3;
Matches 25; Conservative 97; Mismatches 219; Indels 0; Gaps 0;

QY 756 GCCCAGCGGTGATATGCGAGGGGTAAAGCGTTCGATGCTTCATATGCAATGCGGGCCAG 815
DB 423 VVNNLSYNKLSDSVFCPLPKSIQILDNNNQIOTVPKETIHLMALRELNIAFNFLTDLP 482
QY 816 TCGAAACCCGGTGGATAACCGTTACCGGTGTCTGACACCTCCGCCCATCAGATCGCC 875
DB 483 GCSHFSRLSVLNIEMNFILSPSLDFVQSCOEVTNLNRPFRCTCELKNFIOLEYSEV 542
QY 876 ATCATGGGTGATAGGACACAAATGAAATACACAGATTAATTCAGGGAACGTTCTG 935
DB 543 MMVGWSDSYTCYPLNLGRTRLDVHLHLSLNTALLIVIVIMLVGLAVAFCCCLHFD 602
QY 936 GICTTAGCGGTGATGATGTTTGTCTGACAAATAGTAGCGGTGACAGCATATCAGA 995
DB 603 LPWYRLMQCCTQTHRRVKTQEQLRNVRHAFISYSEHDSLVKVNELIPNLEKEDGS 662
QY 996 CGGCTCAGTCTGTATATCTGTCATGCCACTATGCCAGATGCCAGATCAGGATTTA 1055
DB 663 ILICLYESYFDPKSISENIVSFIKSYKSFVLSNPNVONEWCHVEFFFAHNLPHENS 722
QY 1056 AATTCGCCGATATCGTCAAGTCTCAGGATCGAAGGAAG 1096
DB 723 DHIILILEPPIPCIPTRYHKLKALLEKKAYLEWPKDRK 763

RESULT 32

US-10-184-634-414
; Sequence 414, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 414
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-414

Query Match 2.7%; Score 32.4; DB 9; Length 811;
Best Local Similarity 7.3%; Pred. No. 3;
Matches 25; Conservative 97; Mismatches 219; Indels 0; Gaps 0;

QY 756 GCCCAGCGGTGATATGCGAGGGGTAAAGCGTTCGATGCTTCATATGCAATGCGGGCCAG 815
DB 423 VVNNLSYNKLSDSVFCPLPKSIQILDNNNQIOTVPKETIHLMALRELNIAFNFLTDLP 482

us-09-674-277-2.rnpb

[illegible]


```

; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Zhang,Zemin
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-123-155-90

Query Match 2.7%; Score 32.2; DB 9; Length 468;
Best Local Similarity 7.0%; Pred.No.2.4;
Matches 27; Conservative 117; Mismatches 241; Indels 0; Gaps 0;

QY 741 TTGCTGCTGATCAGCCAGCGGTGATGATGAGGTAACGGTAAACGGTCGCATGGCTTCATTA 800
DB 58 YPENPGQNYISFVDMNLLKAITKEKIEKERQSIIRSPIDNKLNVEDVDSTKNRKLI 117
QY 801 TGCATGCGCGCCAGTCGAAACCGGTGGAAATACCGTTACCGGTGCTCTGCACACCTTCC 860
DB 118 DDVDTKSLGDHKFDQDPDLGDLGDTPLTAEDIVHKIARIYEENDRAVFDKIVSKLN 177
QY 861 GCATCATGATCGCCCATCATGCGGTGATAGAGGACACACATGAATACACACATATTC 920
DB 178 LGLITESQATLEDEVAEVLQKLISKANNYEEDPNKPTSWENQAGKIKEKTPHAAIQ 237
QY 921 AGGGAAACGTTCTGCTGCTTACGGGTGATGATGAGTTTGTCTGCAATAGTAGAAGCGG 980
DB 238 DGLAKGENDETIVSNLTLTNGLERRTKTSYSEDNFEELQYFPNFVALLKSIDSEKAKEKE 297
QY 981 TGACAGCATATCAGACGGCTCAGTCCTGCTGCTATATTACTGTGATGCCCACTATGCCAGATG 1040
DB 298 TLTIIMKTLDFVKMMVKYGTISPEEGVSVLENIDEMIALQTNKLEKKNATDNISKLFPA 357
QY 1041 ACCAGATCAGGTTTAAATTCGCCGATAATCCGTCGAACTCTGAGGATGGAAGGAGGTGA 1100
DB 358 PSEKSHETDSTKEAAKMEKEYGLSKDSIKDDNSNPGKTDPEKGTGTEAVLEAIRKNIE 417
QY 1101 AGGCTGTCTCGAAAGGAATAAAG 1125
DB 418 WLKKHKKGNKEDYDLKMRDFINK 442

RESULT 36
US-09-738-626-451/c
; Sequence 451, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18

```

```

; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1734
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_022403
US-09-917-800A-1734

Query Match      2.7%; Score 32.2; DB 10; Length 1689;
Best Local Similarity 52.6%; Pred. No. 5.6;
Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1013 ATTACTGTCAGCCACTATGGCAGATGACAGATCAGGTAAATTCCTCCCGAATATCCG 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 ATCCCTGACATGGCCATTTTGAAAAATCTCAGCAACAGAAATCAAGTTCCCGAGAAATTTG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1073 TCGAAGCTGAGGATGAAGAGGTGAAGCTGTCTCTGAAAGGAATAAAGTGACATC 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 TTTAAACCAAGTTCATAAGCTTGGTGAGTTAATAAAGAGGTGCTGCTGGATTTT 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1133 ATGCCCTCTTTT 1145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 ATTCCTTTGATT 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 38
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match      2.7%; Score 32.2; DB 9; Length 3309400;
Best Local Similarity 45.5%; Pred. No. 3.5e+02;
Matches 115; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 270 TCTTCTGCGGCTGGCTGCCATCGGAAGCGTTCCGTCGGGGATATAAATAATCGGCG 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425031 TGTTCGAGCGGCTCTTCCACCATCTGATGCGGCTTTCGTATCGTCCAGAGTGTGCT 424972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 AGTGGCGGTCATGACAGACATCCCCACAGGGTACAGCGGTCCTGTGCATCTTC 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424971 AAGCGTGCAGCCCGGAGTGGCGGAGTGGCGGTCGACAAGCAATCCCGTCTCCCTTCC 424912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 TGAATGATCAGGATCCCGCGTCTCATCTGGGATACGGGCGGAGCTGAC 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 424911 CGACTGGATGGGAGGCCGCCAACCCCGGCCGCGGATGACCGGTGTGCCGTGCTTGC 424852
QY 450 GCTTCAGCCAGTACCATACCAAGCGCTTCATTTTCCGAAGCATGACCAACACTGGCA 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424851 GCCTCCATGGCGAGGAGTCCGAAGGACTCATTTAAACCTTTGACAGGCCACGATGTCGCC 424792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 510 ATCCGGGTAGACCG 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424791 GCCCGATACACGG 424779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 39
US-09-815-242-4609/c
; Sequence 4609, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4609
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4609

Query Match      2.7%; Score 32; DB 10; Length 1311;
Best Local Similarity 49.4%; Pred. No. 5.6;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 633 AGCCAGCGAATGATTTCCCTTCCATCTTCCATGATACATACACGAGCATATAATCA 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1166 ACCCAGGAATAGTTGCCACTTCATTTTAAATTTGTTTAAACACGACATAGTTCA 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 693 TGTCTTTTTCGGGACGTAGCATCCCGACCTGAACGATAAGCGGAACATTGTCTGCTGAT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1106 TGCTCATTACACCAATAAATCTTGAATTCATCAATTAATGGCATATATGTTTACCGAT 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 753 GCAGCCAGGCGTGATGATGAGGGGTAAACGTCGCGATCGGCTTCATTA 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1046 ACAGTAATGTTGGAAATAAGCAAGAAATGTTTCCAATCTTCATTA 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 40
US-09-815-242-8107/c
; Sequence 8107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

```

```

RESULT 41
US-10-184-644-402/c
; Sequence 402, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECURED AND TR
; TITLE OF INVENTION: ACIDS ENCODING

```

RESULT 42
US-10-184-634-402/c
; Sequence 402, Application US/10184634
; Publication No. US2003006984A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Chennoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRA

;	APPLICANT:	Smith,Victoria
;	APPLICANT:	Stewart,Timothy A.
;	APPLICANT:	Tunas,Daniel
;	APPLICANT:	Watanabe,Colin K
;	APPLICANT:	Wood,William
;	APPLICANT:	Zhang,Zemin
;	TITLE OF INVENTION:	SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;	TITLE OF INVENTION:	ACIDS ENCODING THE SAME
;	FILE REFERENCE:	P333OR1G30
;	CURRENT APPLICATION NUMBER:	US/10/123,155
;	CURRENT FILING DATE:	2002-04-15
;	Prior Application removed -	See Palm or File Wrapper
;	NUMBER OF SEQ ID NOS:	550
;	SEQ ID NO 358	:
;	LENGTH:	1049
;	TYPE:	PRT
;	ORGANISM:	Homo Sapien
;	US-10-123-155-358	:

Query Match	2.7%	Score 31.8;	DB 9;	Length 1049;
Best Local Similarity	4.1%	Pred. No. 5.7;		
Matches 23;	Conservative 151;	Mismatches 391;	Indels 0;	Gaps 0;

QY	493	TGACCACCACTGGCAATCCCGTAGACCGGTAAACGCTGGGAAGGCACCTGCCATT	552
Dd	790	HGAG..YNDD...KM.NN.BN...SAM..K..MNSS.H.RY...TH...RMA.SY..NNM	731
QY	553	ACACACTCGCTCATCCAGGTGTTCTGTCGCTGACGACGACGCTTCTGTAATCTT	612
Dd	730	..NM.SYSGNS.Y..BAA..NDS.H.A..NM.G..MMWS.SM.CNMT.S.NM.N..KCH	671
QY	613	CAGCGCCGGCGCCACACGACGACGAGCAGGAATGATTTCCCTTCATCTTCAGCTGATACA	672
Dd	670	.BCS....S.SNMS.H....M..NM....RYNHC.YW.BH.DNCY...AY.S.S.KAYSAS	611
QY	673	ATACACGACGATAAATTCATGTCCTTTTCGGGACGTAGCATCCCCACCTGAACGATAA	732
Dd	610	SS..HNHNKK.M..BM.NMA.N.KDA.C.S..RDSNSS.H.B..MD....TASD..H.YNN	551
QY	733	CGCGAACATGTCTGCTGATGCAGCCCAGCGGTGGATATCAGGGGTAAACGTCGCATGG	792
Dd	550	S.H.RY.T.....SCN.A.S..NCS.N.GM..S.D..HSSMB....SNMS.H.A.CRMRG	491
QY	793	CTTCATATGCAATGGGGCGCAGTCGAAACCCGGTGGAATAACCGTATACCGGTGTCTGA	852
Dd	490	S.NBSK.ST.MNM.YGSYTRMHY.RD....B...RS.BSAYTNG.CB.SSHCS.S.MN	431
QY	853	CACCTTCGCCATCAGATGCGCCATCATCGGTGAGATAGGACACAAATGAATCACACA	912
Dd	430	BS..H.BM.YM..M.KS.NT.M..NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.S	371
QY	913	GATAAATTCAGGGAACGCTTCTGGCTTACGGGTGATGAGTTTTTTTCTGCACAATAG	972
Dd	370	M.SS.T.S..NKSTYRB....N.S.H....S..D..MTHC..MT..N.S.H....NM	311
QY	973	TGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCTATATTACTGTCTATGCCACTAT	1032
Dd	310	WY..BD..SNSD.Y.BM.A.TH.TNB.....SNNMG.TG....TNRGY.GNCS.H....	251
QY	1033	GGCAGATGACCATCAGGTTTAAA	1057
Dd	250	N.NN.HH...MT.KNNR.R..A.A	226

RESULT 44

US-09-891-641-35/c

; Sequence 35, Application US/09891641

; Patent No. US20020155612A1

; GENERAL INFORMATION:

; APPLICANT: Ye, Rick

; APPLICANT: Bedzyk, Laura

; APPLICANT: Wang, Tao

; TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN BACILLUS SPECIES

```
; FILE REFERENCE: C11686 US NA
; CURRENT APPLICATION NUMBER: US/09/891.641
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft 97
; SEQ ID NO 35
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-891-641-35

Query Match      2.7%; Score 31.8; DB 9; Length 1146;
Best Local Similarity 45.7%; Pred. No. 6;
Matches 111; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 272 TTCTCTGGCGTGGCTGCCATCATCCGGAAGCGTTCGGGTCCGGATATAAAATTCGGCGAG 331
Db 1050 TTTTGGCGCTTCTGTCCTCCAGCTTTCCGCGAGAGCTCCGGCTTATGGTAAAGCTGTTCAT 991
QY 332 TGGCGCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTCACTTCTCTG 391
Db 990 CCGCGGGGCAAAAGCAGCACTGTCCACGATTTTGATCAGAAACCGGTTTCTCGCTCGCG 931
QY 392 AATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAACGGGCGACGCGGAGACTGACGC 451
Db 930 GATGATTTCCGATGCCGCGGATATCTGTGCGCATCGCGGTTGTTCTCGCGCCATTC 871
QY 452 TTCAGCCAGTACCATAACAAAGCGTTCATTTTCCGAAGGCATGACCAACACACTGGCAAT 511
Db 870 CTCAGCACATTCATACGAGGCGCTTCTTAATGCTGATGCGACAGATACATCCGCAAG 811
QY 512 CCG 514
Db 810 CTG 808

RESULT 45
US-09-759-130B-343/c
; Sequence 343, Application US/09/759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: US.
; FILE REFERENCE: MPI00-5350NNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
```

```
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 343
; LENGTH: 3153
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(3153)
; OTHER INFORMATION: n - A,T,C or G
US-09-759-130B-343

Query Match      2.7%; Score 31.8; DB 9; Length 3153;
Best Local Similarity 51.0%; Pred. No. 11;
Matches 75; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 396 ACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGCGACCGGAGACTGACGGTTCA 455
Db 2000 ACTCCACCGTGGCTGGCACTGTGCTGGCGAGCACTGGCTGGCCTGCACTGAGGTGCT 1941
QY 456 GCCAGTACCATAACCAACGCTTCATTTCCGAAGGCATGACCAACCACTGGCAATCCGG 515
Db 1940 GCCAGGACAGTTCTTCCAGACTTCTCTTGAGGGGGTCTCCAGCTCCCTGGGACCCACA 1881
QY 516 TAGACCGGTACCGCTGGGAAAGGGCA 542
Db 1880 GGGGCCCATGTAGCTGGAAGTAGGGAA 1854

RESULT 46
US-10-189-123-73/c
; Sequence 73, Application US/10189123
; Publication No. US20030082586A1
; GENERAL INFORMATION:
; APPLICANT: KIRST, Susan J.
; APPLICANT: HOLTZMAN, Douglas A.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; APPLICANT: BARNES, Thomas S.
; TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 10147-11U3
; CURRENT APPLICATION NUMBER: US/10/189,123
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 3153
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3043)..(3043)
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3048)..(3048)
; OTHER INFORMATION: unknown
US-10-189-123-73

Query Match      2.7%; Score 31.8; DB 9; Length 3153;
Best Local Similarity 51.0%; Pred. No. 11;
Matches 75; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
```



```
Db 222 CC..CYV.TYBTMM..A.H.HSAM.S.SSS.SN....S.SBST.H.HSSTMTYNSBSKM 163
Qy 817 CGAACCCGGTGAATAACGGTTACC 842
Db 162 .T.AMYM.CSNHSSMHSHS.KYHS 137

RESULT 49
US-09-381-624A-5/c
; Sequence 5, Application US/09381624A
; Patent No.. US20020068349A1
; GENERAL INFORMATION:
; APPLICANT: Horinouchi, Sueharu
; APPLICANT: Saitoh, Kohki
; APPLICANT: Takahashi, Eisaku
; TITLE OF INVENTION: GENE ENCODING TREHALOSE PHOSPHORYLASE, VECTOR CONTAINING THE GENE
; TITLE OF INVENTION: TRANSFORMANT TRANSFORMED BY THE GENE, AND METHOD FOR PRODUCING R
; TITLE OF INVENTION: TREHALOSE PHOSPHORYLASE WITH THE USE OF TRANSFORMANT
; FILE REFERENCE: FJIN-100
; CURRENT APPLICATION NUMBER: US/09/381,624A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 09/381,624
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 2199
; TYPE: DNA
; ORGANISM: Grifola frondosa
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(3)
; NAME/KEY: terminator
; LOCATION: (2197)..(2199)
US-09-381-624A-5

Query Match 2.7%; Score 31.6; DB 10; Length 2199;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 103; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 290 CATCATCCGGAAGCGTTCCGGTCGGGATATAAAATCGGCAGTGCAGCGGTCATGCAGA 349
Db 1986 CATCGTGTATACAGGTCCTCGTGTACAAAGTCAGCATGTGCTGCGCAACTGCTGC 1927
Qy 350 CACATCCCCACGGGTAAACAGCTCCCTGTACATTTCTTGATGACATCAGGATCCC 409
Db 1926 GTTGTCGCGCGCTCGCAGAGGTATCCGCTCTTTCCATGCTCAATCTGCAGCGGATGCC 1867
Qy 410 GCCCGTCTCACTGGCGATAACGGCACGCCGAGACTGACGCTTCAGCCAGTACCATACC 469
Db 1866 GCCCGTGGCATGCGATGAGCGGGCTTGCCGTGGCAAGCCCTCTGAACCTTTACTTC 1807
Qy 470 AAACGCTTCATTTTCCGAAGGCATGACCCACACACTGGCAAT 511
Db 1806 GAAGCCCTCACCGCTCGAGAGCTGTAGCGCGAACTTCGCATT 1765

RESULT 50
US-09-381-624A-4/c
; Sequence 4, Application US/09381624A
; Patent No.. US20020068349A1
; GENERAL INFORMATION:
; APPLICANT: Horinouchi, Sueharu
; APPLICANT: Saitoh, Kohki
; APPLICANT: Takahashi, Eisaku
; TITLE OF INVENTION: GENE ENCODING TREHALOSE PHOSPHORYLASE, VECTOR CONTAINING THE GENE
; TITLE OF INVENTION: TRANSFORMANT TRANSFORMED BY THE GENE, AND METHOD FOR PRODUCING R
; TITLE OF INVENTION: TREHALOSE PHOSPHORYLASE WITH THE USE OF TRANSFORMANT
; FILE REFERENCE: FJIN-100
; CURRENT APPLICATION NUMBER: US/09/381,624A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 09/381,624
; PRIOR FILING DATE: 1999-12-15
```

```
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2817
; TYPE: DNA
; ORGANISM: Grifola frondosa
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(65)
; NAME/KEY: Intron
; LOCATION: (177)..(239)
; NAME/KEY: Intron
; LOCATION: (486)..(547)
; NAME/KEY: Intron
; LOCATION: (1044)..(1160)
; NAME/KEY: Intron
; LOCATION: (1175)..(1228)
; NAME/KEY: Intron
; LOCATION: (1282)..(1350)
; NAME/KEY: Intron
; LOCATION: (1551)..(1618)
; NAME/KEY: Intron
; LOCATION: (1872)..(1928)
; NAME/KEY: Intron
; LOCATION: (2300)..(2356)
; NAME/KEY: terminator
; LOCATION: (2809)..(2811)
; NAME/KEY: terminator
; LOCATION: (2815)..(2817)
US-09-381-624A-4

Query Match 2.7%; Score 31.6; DB 10; Length 2817;
Best Local Similarity 46.4%; Pred. No. 13;
Matches 103; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 290 CATCATCCGGAAGCGTTCCGGTCGGGATATAAAATCGGCAGTGCAGCGGTCATGCAGA 349
Db 2598 CATCGTGTATACAGGTCCTCGTGTACAAAGTCAGCATGTGCTGCGCAACTGCTGC 2539
Qy 350 CACATCCCCACGGGTAAACAGGTCCTGTACATTTCTTGATGACATCAGGATCCC 409
Db 2538 GTTGTCGCGCGCTCGCAGAGGTATCCGCTCTTTCCATGCTCAATCTGCAGCGGATGCC 2479
Qy 410 GCCCGTCTCACTGGCGATAACGGCACGCCGAGACTGACGCTTCAGCCAGTACCATACC 469
Db 2478 GCCCGTGGCATGCGATGACGGGCTTGCTGCGTGCAGAGGCTCTGAACCTTTACTTC 2419
Qy 470 AAACGCTTCATTTTCCGAAGGCATGACCCACACACTGGCAAT 511
Db 2418 GAAGCCCTCACCGCTCGAGAGCTGTAGCGCGAACTTCGCATT 2377

Search completed: June 19, 2003, 06:35:47
Job time : 311.567 secs
```